KUNS-1929, OHSTPY-HEP-T-04-008, MCTP-04-13

Searching for realistic 4d string models with a Pati-Salam symmetry

– Orbifold grand unified theories from heterotic string compactification on a $\mathbb{Z}_6$ orbifold

Tatsuo Kobayashi

Department of Physics, Kyoto University, Kyoto 606-8502, Japan.

Stuart Raby

Department of Physics, The Ohio State University, Columbus, OH 43210 USA.

Ren-Jie Zhang

Michigan Center for Theoretical Physics, Randall Laboratory,

University of Michigan, Ann Arbor, MI 48109 USA.[*]

(Dated: September 8, 2004)

Abstract

Motivated by orbifold grand unified theories, we construct a class of three-family Pati-Salam models in a $\mathbb{Z}_6$ abelian symmetric orbifold with two discrete Wilson lines. These models have marked differences from previously-constructed three-family models in prime-order orbifolds. In the limit where one of the six compactified dimensions (which lies in a $\mathbb{Z}_2$ sub-orbifold) is large compared to the string length scale, our models reproduce the supersymmetry and gauge symmetry breaking pattern of 5d orbifold grand unified theories on an $\mathrm{S}^1/\mathbb{Z}_2$ orbicircle. We find a horizontal $2+1$ splitting in the chiral matter spectra – 2 families of matter are localized on the $\mathbb{Z}_2$ orbifold fixed points, and 1 family propagates in the 5d bulk – and identify them as the first-two and third families. Remarkably, the first two families enjoy a non-abelian dihedral D_4 family symmetry, due to the geometric setup of the compactified space. In all our models there are always some color triplets, i.e. $(\mathbf{6},\mathbf{1},\mathbf{1})$ representations of the Pati-Salam group, survive orbifold projections. They could be utilized to spontaneously break the Pati-Salam symmetry to that of the Standard Model. One model, with a 5d E_6 symmetry, may give rise to interesting low energy phenomenology. We study gauge coupling unification, allowed Yukawa couplings and some of their phenomenological consequences. The E_6 model has a renormalizable Yukawa coupling only for the third family. It predicts a gauge-Yukawa unification relation at the 5d compactification scale, and is capable of generating reasonable quark/lepton masses and mixings. Potential problems are also addressed, they may point to the direction for refining our models.

PACS numbers: 11.25.Mj,12.10.-g,11.25.Wx

[*]Also at Institute of Theoretical Physics, The Chinese Academy of Sciences, Beijing 100080, China.

I. INTRODUCTION AND MOTIVATION

The Standard Model (SM) has been a cornerstone of modern-day particle physics. Although during the past three decades it has passed all experimental tests, nevertheless there are many open questions remaining to be answered. We have yet to understand — (i) the mechanism of electro-weak symmetry breaking, and find the Higgs boson which might be responsible for this breaking; (ii) the quantized fermion charges (why the up and down quarks have charges $2/3$ and $-1/3$ respectively) and the weak mixing angle (why it is 0.23); (iii) the 3 replicas of quarks and leptons, the observed fermion mass hierarchy, and the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix and its leptonic cousin. We ought to understand these problems from more fundamental principles, rather than simply take various charges, masses, mixing angles and CP phases as input parameters as in the SM. Addressing these questions may eventually lead us to a more fundamental theory such as string theory at high energy scales.

String theory [1] is a leading candidate for a consistent theory of quantum gravity. It has a rich structure and many believe it can easily accommodate the SM as a subset. Moreover there have been many attempts in the past to construct supersymmetric generalizations of the SM (which will be loosely referred to as the minimal supersymmetric standard model, or the MSSM) or grand unified theories (GUTs) from the heterotic string [2, 3, 4, 5, 6] and superstrings [7]. Partially successful results have been obtained. For example, many string theoretical models can explain the existence of three chiral families at low energy scales [7, 8, 9, 10, 11, 12, 13, 14], and in principle can also provide a natural framework for understanding the fermion masses and mixings [15]. In this paper, we construct a new class of three-family models in the heterotic string theory.

Before presenting our models, it is important to note some caveats common to all known string models. They are due to two main difficulties facing the string theoretical model constructions. The first difficulty concerns the compactification of the string itself, i.e. the mechanism by which the desirable string vacuum is selected. The vacua of string theory compactifications are parameterized by many scalar fields with flat potentials. These fields are the modulus fields. They characterize the sizes and shapes of the compactified spaces and the strengths of the string interactions; none of them can be fixed in perturbation theory [16].[1] The modulus problem and the related issue of supersymmetry breaking will not be dealt with in this article, instead we will simply assume that the moduli are fixed by some unknown mechanism at the string scale.

[1] Stabilizing moduli by fluxes in the context of heterotic string theory has been discussed recently in [17].

The second difficulty concerns our ignorance of the physics between the electro-weak and unification scales. Except for some indication that the SM gauge couplings may unify at about 10^{16} GeV in certain supersymmetric extensions of the SM with minimal matter content [18], we can hardly have any confidence in extrapolating the low energy data by some 14 to 16 orders of magnitude to the unification scale and infer what the gauge symmetries, matter spectra and physical parameters are at that scale. Hence, we are far from having a clear-cut field theoretical model at the unification scale to which a string-derived model is supposed to match. Any string construction must combine both bottom-up and top-down analyses.

The new class of string models in this paper are mainly motivated by the recent discussions on orbifold GUT models [19, 20]. These GUT models utilize properties of higher-dimensional field theories, and have some advantages over conventional 4d GUTs. For example, GUT symmetry breaking can be accomplished by an orbifold parity, instead of by a complicated Higgs sector. The doublet-triplet splitting problem, which plagues conventional GUTs, can also be solved by assigning appropriate orbifold parities to the doublet and triplet Higgs bosons. Note, however, that like all field theoretical models in higher dimensions, these GUT models are not renormalizable quantum field theories. They can only make sense as low-energy effective theories of some more fundamental theory with better ultra-violet (UV) behavior. Our string models provide exactly such kind of UV completions, in the sense that they reproduce many interesting features of the orbifold GUTs in certain low energy limits. (Connections between orbifold GUTs and an SO_{10} string model have already been established in ref. [21]. The present paper contains more detailed discussions on these connections. See also the recent paper [22].)

To make the connections between string and field theoretical models more concrete, we consider some examples, in particular, the 5d SO_{10} model of ref. [20] and a generalization with bulk gauge group E_6. In these models, the extra dimension is taken to be an orbicircle $\mathrm{S}^1/\mathbb{Z}_2$ and the 4d effective theory has a Pati-Salam (PS) symmetry, $\mathrm{SU}_{4\mathrm{C}}\times\mathrm{SU}_{2\mathrm{L}}\times\mathrm{SU}_{2\mathrm{R}}$ [23]. The technical apparatus we adopt to build string models is the simplest abelian symmetric orbifold compactification [4, 6, 9, 10, 11, 24] of the heterotic string [2]. More specifically, we consider a non-prime-order $\mathbb{Z}_6$ orbifold (or equivalently, $\mathbb{Z}_2 \times \mathbb{Z}_3$) model with the orbifold twist vector $\mathbf{v}_6 = \frac{1}{6}(1, 2, -3)$. To achieve three chiral PS families at low energies, we also introduce several (in fact, two) discrete Wilson lines [25]. [2] It is obvious that the third compactified complex dimension has a $\mathbb{Z}_2$ symmetry in the $\mathbb{Z}_6$

[2] Prime-order orbifold models (such as the $\mathbb{Z}_3$ orbifold models) with Wilson lines [9, 10, 11] and non-prime-order orbifold models without Wilson lines [24] have been extensively studied in the literature. Non-prime-order orbifold models with Wilson lines, on the other hand, possess a number of complications, and to our knowledge they

model, hence it can consistently be taken to be the root lattice of the SO_4 Lie algebra. The string models are effectively 5d when the length of one of the SO_4 simple roots is large compared to the string scale, while all other dimensions are kept comparable to the string scale (i.e. the geometry of the compactified space is equivalent to that of the orbifold GUTs, $S^1/\mathbb{Z}_2$). In this limit, the $\mathbb{Z}_6$ heterotic models are similar to the orbifold GUT models in the following respects:

- The 5d $\mathcal{N} = 2$ supersymmetry[3] is broken to that of $\mathcal{N} = 1$ in 4d by the $\mathbb{Z}_2$ orbifold twist and the "bulk" gauge group is broken to two different regular subgroups at the two inequivalent fixed points by degree-2 non-trivial gauge embedding and Wilson line. The surviving gauge group in the 4d effective theory is the intersection of groups at the fixed points. It is the PS group in our models. More specifically, we find two types of models. In the first type we have an E_6 symmetry in the 5d bulk which is broken to SO_{10} and $SU_6 \times SU_2$ respectively. In the second type we have an $SO_{10} \times SU_2$ in the bulk, broken to PS at one of the two fixed points.

- Untwisted-sector and twisted-sector states that are not localized on the $\mathbb{Z}_2$ fixed points of the SO_4 lattice can be identified with the "bulk" states of the orbifold GUT. Interpretation of the Kaluza-Klein (KK) towers of the bulk gauge and matter fields agree in the string-based and orbifold GUT models.

- Twisted-sector states that are localized on the $\mathbb{Z}_2$ fixed points of the SO_4 lattice have no field theoretical counterparts, although they can correctly be identified with the "brane" states of the orbifold GUT. In the orbifold GUT models, these states are only constrained by the requirement of (chiral) anomaly cancellation.

Of course, string theoretical models are more intricate than the corresponding field theoretic orbifold GUT models. They need to satisfy more stringent consistency conditions and thus they are physically more constrained. We find it is highly non-trivial (or impossible) to implement all the features of the orbifold GUTs. For example, we cannot arbitrarily place the three families of quarks and leptons in the bulk or on either brane. Moreover, the very act of obtaining three families, along with their respective locations, is fixed by the requirement that the gauge embeddings and

have not been studied to the same extent. Our work can be regarded as the first serious attempt at constructing three-family models from non-prime-order orbifolds.

[3] By $\mathcal{N} = 2$ supersymmetry in 5 or 6d, we mean the minimal number of supersymmetries in these dimensions, (i.e. the fermions satisfy the pseudo-reality condition). It reduces to $\mathcal{N} = 2$ in 4d by dimensional reduction and is sometimes called $\mathcal{N} = 1$ supersymmetry in the literature.

Wilson lines have to satisfy the modular invariance conditions [4, 26]. In addition, we cannot utilize the orbifold projections to remove all the $(\mathbf{6},\mathbf{1},\mathbf{1})$ color-triplet states as in the SO_{10} orbifold GUTs [20] and at the same time obtain three families. We also find many massless states carrying unconventional representations under the SM gauge group. These exotic states are commonplace in almost all known three-family models. Whether these models can give rise to satisfactory phenomenology needs more detailed knowledge of the low-energy effective actions. The present status of our analysis is contained in this paper.

The paper is organized as follows. In sect. II we briefly review 5d field theories on the orbicircle $\mathrm{S}^1/\mathbb{Z}_2$ and present two orbifold GUT models with bulk gauge groups SO_{10} and E_6. The latter (model A1[4]) is a novel 5d model with many nice phenomenological features. Then in sect. III we discuss the heterotic string construction of model A1. Using this model as a guide we compare the heterotic string construction with generic orbifold GUT models by restricting the compactified space to a specific type (which is referred to as the orbifold GUT limit). We show the equivalence between the matter states (in the untwisted and some twisted sectors) in string-based models and the bulk states in orbifold GUTs, as well as their KK excitations. We interpret orbifold parities (for the bulk states) in the orbifold GUTs in string theory language, and explain why the gauge embeddings and Wilson lines cannot project away all the $(\mathbf{6},\mathbf{1},\mathbf{1})$ color-triplet states. These states may be needed to break the PS group to that of the SM, as in the field theoretical model of sect. II B. In sect. IV we focus on more of the phenomenological aspects of model A1. In sect. IV A we discuss gauge coupling unification and the determination of the compactification and string scales. In sect. IV B we examine the allowed Yukawa couplings (at both the renormalizable and non-renormalizable levels) and their phenomenological consequences, concentrating on the possibility of breaking the PS symmetry, mass generation for the color-triplet fields and SM fermions, and proton stability. We conclude in sect. V, listing the pros and cons of the present models. Hopefully one can learn from the problems to design better models in the future.

We have made an effort to make the paper more accessible to field theory model builders. Many of the details of string constructions are relegated to four appendices. In appendix A we review the construction of non-prime-order orbifold models with Wilson lines, highlighting its differences with the prime-order orbifold construction. In appendix B 1 we present three three-family $\mathbb{Z}_6$ models with PS gauge symmetry. The complete matter spectra are listed in appendix B 2, where we also explain the notation for the twisted-sector states. In appendix C 1 we review the string

[4] This model is denoted A1 since it corresponds to the first of several string models discussed in the paper.

selection rules necessary for determining non-trivial Yukawa couplings in a 4d effective theory, and in appendix C 2 list some allowed couplings involving operators of interest in model A1. Finally, in appendix D we study gauge coupling unification and derive the Georgi-Quinn-Weinberg (GQW) relations [27] in the orbifold GUT limit. These relations allow us to determine various mass scales in our models.

II. 5D ORBIFOLD GUT MODELS ON $S^1/\mathbb{Z}_2$

Let us briefly review the geometric picture of orbifold GUT models compactified on an orbicircle $S^1/\mathbb{Z}_2$. The space group of $S^1/\mathbb{Z}_2$ is composed of two actions, a translation, $\mathcal{T}: x^5 \to x^5 + 2\pi R$, and a space reversal, $\mathcal{P}: x^5 \to -x^5$. There are two (conjugacy) classes of fixed points, $x^5 = 2n\pi R$ and $(2n+1)\pi R$, where $n \in \mathbb{Z}$.

The space group multiplication rules imply $\mathcal{TPT} = \mathcal{P}$, so we can replace the translation by a composite $\mathbb{Z}_2$ action $\mathcal{P}' = \mathcal{PT} : x^5 \to -x^5 + 2\pi R$. The orbicircle $S^1/\mathbb{Z}_2$ is equivalent to an $\mathbb{R}/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ orbifold, whose fundamental domain is the interval $[0, \pi R]$, and the two ends $x^5 = 0$ and $x^5 = \pi R$ are fixed points of the $\mathbb{Z}_2$ and $\mathbb{Z}_2'$ actions respectively.

A generic 5d field Φ has the following transformation properties under the $\mathbb{Z}_2$ and $\mathbb{Z}_2'$ orbifoldings (the 4d space-time coordinates are suppressed),

$$\mathcal{P}: \Phi(x^5) \to \Phi(-x^5) = P\Phi(x^5)\,, \qquad \mathcal{P}': \Phi(x^5) \to \Phi(-x^5 + 2\pi R) = P'\Phi(x^5)\,, \tag{2.1}$$

where P, $P' = \pm$ are *orbifold parities*. In general cases $P' \neq P$; this corresponds to the translation $\mathcal{T}$ being realized non-trivially by a degree-2 Wilson line (i.e., background gauge field). The four combinations of orbifold parities give four types of states, with wavefunctions $\Phi_{++}(x^5) \sim \cos(mx^5/R)$, $\Phi_{+-}(x^5) \sim \cos[(2m+1)x^5/2R]$, $\Phi_{-+}(x^5) \sim \sin[(2m+1)x^5/2R]$ and $\Phi_{--}(x^5) \sim \sin[(m+1)x^5/R]$, where $m \in \mathbb{Z}$. The corresponding KK towers have masses

$$M_{\rm KK} = \begin{cases} m/R & \text{for } (PP') = (++)\,, \\ (2m+1)/2R & \text{for } (PP') = (+-) \text{ and } (-+)\,, \\ (m+1)/R & \text{for } (PP') = (--)\,. \end{cases} \tag{2.2}$$

Note that only the Φ_{++} field possesses a massless zero mode.

A. An SO_{10} orbifold GUT

Consider the 5d orbifold GUT model of ref. [20]. The model has an SO_{10} symmetry broken to the PS gauge group, $SU_{4C} \times SU_{2L} \times SU_{2R}$, in 4d, by orbifold parities. The compactification scale

$M_c = (\pi R)^{-1}$ is assumed to be much less than the cutoff scale. (In string theory the cutoff scale is given by the string scale $M_{\rm string}$.)

The gauge field is a 5d vector multiplet $\mathcal{V} = (A_M, \lambda, \lambda', \sigma)$, where A_M, σ (and their fermionic partners λ, λ') are in the adjoint representation (**45**) of SO_{10}. This multiplet consists of one 4d $\mathcal{N} = 1$ supersymmetric vector multiplet $V = (A_\mu, \lambda)$ and one 4d chiral multiplet $\Sigma = ((\sigma + iA_5)/\sqrt{2}, \lambda')$. We also add a 5d hypermultiplet $\mathcal{H} = (\phi, \phi^c, \psi, \psi^c)$ in the **10** representation. It decomposes into two 4d chiral multiplets $H = (\phi, \psi)$ and $H^c = (\phi^c, \psi^c)$ in complex conjugate representations. This model has an $\mathcal{N} = 2$ extended supersymmetry. The 5d gravitino $\Psi_M = (\psi^1_M, \psi^2_M)$ decomposes into two 4d gravitini ψ^1_μ, ψ^2_μ and two dilatini ψ^1_5, ψ^2_5. To be consistent with the 5d supersymmetry transformations one can assign positive parities to $\psi^1_\mu + \psi^2_\mu$, $\psi^1_5 - \psi^2_5$ and negative parities to $\psi^1_\mu - \psi^2_\mu$, $\psi^1_5 + \psi^2_5$; this assignment partially breaks $\mathcal{N} = 2$ to $\mathcal{N} = 1$ in 4d.

The orbifold parities for various states in the vector and hyper multiplets are chosen as follows [20] (where we have decomposed all the fields into PS irreducible representations)

States	P	P'	States	P	P'
$V(\mathbf{15},\mathbf{1},\mathbf{1})$	$+$	$+$	$\Sigma(\mathbf{15},\mathbf{1},\mathbf{1})$	$-$	$-$
$V(\mathbf{1},\mathbf{3},\mathbf{1})$	$+$	$+$	$\Sigma(\mathbf{1},\mathbf{3},\mathbf{1})$	$-$	$-$
$V(\mathbf{1},\mathbf{1},\mathbf{3})$	$+$	$+$	$\Sigma(\mathbf{1},\mathbf{1},\mathbf{3})$	$-$	$-$
$V(\mathbf{6},\mathbf{2},\mathbf{2})$	$+$	$-$	$\Sigma(\mathbf{6},\mathbf{2},\mathbf{2})$	$-$	$+$
$H(\mathbf{6},\mathbf{1},\mathbf{1})$	$+$	$-$	$H^c(\mathbf{6},\mathbf{1},\mathbf{1})$	$-$	$+$
$H(\mathbf{1},\mathbf{2},\mathbf{2})$	$+$	$+$	$H^c(\mathbf{1},\mathbf{2},\mathbf{2})$	$-$	$-$

. (2.3)

We see the fields supported at the orbifold fixed points $x^5 = 0$ and πR have parities $P = +$ and $P' = +$ respectively. They form complete representations under the SO_{10} and PS groups; the corresponding fixed points are called SO_{10} and PS "branes." In a 4d effective theory one would integrate out all the massive states, leaving only massless modes of the $P = P' = +$ states. With the above choices of orbifold parities, the PS gauge fields and the $H(\mathbf{1},\mathbf{2},\mathbf{2})$ chiral multiplet are the only surviving states in 4d. The $H(\mathbf{6},\mathbf{1},\mathbf{1})$ and $H^c(\mathbf{6},\mathbf{1},\mathbf{1})$ color-triplet states are projected out, solving the doublet-triplet splitting problem that plagues conventional 4d GUTs.

B. An E_6 orbifold GUT

We now consider a novel 5d orbifold GUT with an E_6 gauge symmetry. In analogy to the model in sect. II A we take the 5d gauge field, given by (V, Σ), in the adjoint representation (**78**) of E_6. In addition to this we add a matter hypermultiplet $H(\mathbf{27}) + H^c(\overline{\mathbf{27}})$.

We define two orbifold parities

$$P = \exp(\pi \mathrm{i} Q_Z/3) \times P_F\,, \qquad P' = \exp[3\pi \mathrm{i}(B-L)/2] \times P'_F\,, \tag{2.4}$$

which break the E_6 via P to SO_{10} and then via P' to PS. Q_Z is the abelian charge in E_6 commuting with SO_{10}, normalized such that the $\mathbf{27}$ decomposes to $\mathbf{16}_1+\mathbf{10}_{-2}+\mathbf{1}_4$, and P_F, P'_F are appropriate discrete flavor charges. (For explicit definition of the parities in the corresponding string model, see sect. III B.) It is easy to obtain the following projections to $(++)$ modes, where the first step follows from P alone and the second follows from the subsequent action of P',

$$\begin{aligned} V = \mathbf{78} \to \quad & \mathbf{45} \to \quad \text{adjoint of PS}\,, \\ \Sigma = \mathbf{78} \to \quad & \mathbf{16}+\overline{\mathbf{16}} \to \quad f_3^c + \overline{\chi}^c\,, \\ \mathbf{27} \to \quad & \mathbf{16} \to \quad f_3\,, \\ \overline{\mathbf{27}} \to \quad & \mathbf{10} \to \quad h\,. \end{aligned} \tag{2.5}$$

In this equation, we have identified the third family of quarks and leptons as well as the MSSM Higgs-doublet pair ($h = H_U + H_D$ where H_U and H_D are the MSSM Higgs doublets responsible for the up- and down-type quark/charged lepton masses),

$$f_3^c = (\overline{\mathbf{4}}, \mathbf{1}, \mathbf{2}), \qquad f_3 = (\mathbf{4}, \mathbf{2}, \mathbf{1}), \qquad h = (\mathbf{1}, \mathbf{2}, \mathbf{2}). \tag{2.6}$$

As a consequence of the fact that the third family and Higgs doublet come from the bulk gauge and $\mathbf{27}$ hypermultiplets we obtain a gauge-Yukawa unification relation,

$$\lambda_t = \lambda_b = \lambda_\tau = g_{4\mathrm{d}} \equiv \sqrt{4\pi\alpha_{\mathrm{GUT}}}\,, \tag{2.7}$$

where $g_{4\mathrm{d}}$ is the 4d gauge coupling constant at the compactification scale. This relation can be seen by inspecting the 5d bulk gauge interaction

$$\int_0^{\pi R} dx^5 \left(g_{5\mathrm{d}} H^c \Sigma H \right) \to g_{4\mathrm{d}}\, h f_3^c f_3\,, \tag{2.8}$$

where $g_{4\mathrm{d}} = g_{5\mathrm{d}}\sqrt{M_c}$.

Of course, we then need to spontaneously break PS to the SM via the standard Higgs mechanism. This can be accomplished when the "right-handed neutrino" fields in

$$\chi^c = (\overline{\mathbf{4}}, \mathbf{1}, \mathbf{2}), \qquad \overline{\chi}^c = (\mathbf{4}, \mathbf{1}, \mathbf{2}) \tag{2.9}$$

obtain non-vanishing vacuum expectation values (vevs)

$$\langle \nu^c \rangle_{\chi^c} = \langle \overline{\nu}^c \rangle_{\overline{\chi}^c} = M_{\mathrm{PS}}. \tag{2.10}$$

We already have one such state but we need more (if only for anomaly cancellation). Consider the addition of three more **27** hypermultiplets given by $3 \times (\mathbf{27} + \overline{\mathbf{27}})$. Upon applying the orbifold parities we find

$$3 \times (\mathbf{27} + \overline{\mathbf{27}}) \to 2(\mathbf{16}) + \overline{\mathbf{16}} + 3(\mathbf{10}) \to 2(\chi^c) + \overline{\chi}^c + 3(C), \tag{2.11}$$

where $C = (\mathbf{6}, \mathbf{1}, \mathbf{1})$. We now have a total $2(\chi^c + \overline{\chi}^c)$ fields. Note, with one C, one χ^c, $\bar{\chi}^c$ pair and a superpotential given by

$$\mathcal{W} = \chi^c \chi^c C + \overline{\chi}^c \overline{\chi}^c C, \tag{2.12}$$

we can give mass to the color triplets and also break PS to the SM along a D- and F-flat direction. (The D-flatness condition requires $\langle \nu^c \rangle_{\chi^c} = \langle \overline{\nu}^c \rangle_{\overline{\chi}^c}$ and $\langle \overline{D}_1 \rangle_{\chi^c} = \langle \overline{D}_1^c \rangle_{\overline{\chi}^c}$, then the F-flatness condition requires further that one of these vevs, say the second, be zero.) In the end, however, we must guarantee that the extra χ^c, $\bar{\chi}^c$, C states obtain mass above the PS breaking scale.

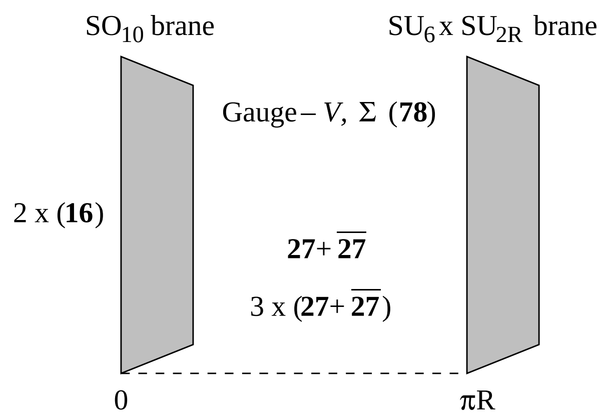

FIG. 1: 5d E_6 orbifold GUT model with bulk and brane states. The bulk gauge symmetry is broken to SO_{10} on the end of world brane at $x^5 = 0$ and to $SU_6 \times SU_{2R}$ at $x^5 = \pi R$. The massless sector of the 4d effective theory has a PS gauge symmetry. In addition, the bulk contains four hypermultiplets, and the SO_{10} brane contains two spinor representations, giving rise to the first two matter families.

But what about the first two families? When constructing an orbifold GUT, one has the option of whether to place the first two families in the bulk or on either brane. One of the main considerations is to avoid rapid proton decay due to gauge exchange and another is to generate a hierarchy of fermion masses. If the compactification scale is much smaller than the GUT scale, say $M_c \ll M_{\rm GUT}$, then it is not possible to place the first two families on the SO_{10} brane. It would however be fine to place them in the bulk or on the $SU_6 \times SU_{2R}$ brane, since in the first case the families are in irreducible representations with massive KK modes, while in the latter case one

family is contained in two irreducible representations $(\mathbf{15},\mathbf{1})+(\overline{\mathbf{6}},\mathbf{2})$, also with massive KK modes. In both cases, gauge exchange takes massless quarks and leptons into massive states. Hence there is *no* problem with proton decay. If however $M_c \geq M_{\mathrm{GUT}}$ then one can place the first two families on either brane. Unfortunately, in string theory, we do not get to choose *easily* where to place the families. It is determined by the choice of vacuum. In the heterotic string version of the model (model A1 in appendix B) we find two families sitting on the SO_{10} brane, as in fig. 1.

III. HETEROTIC STRING CONSTRUCTION OF EFFECTIVE ORBIFOLD GUTS

In appendix A we review the rules for constructing heterotic string models compactified on an abelian symmetric orbifold with discrete Wilson lines. Then in appendix B we construct three three-family $\mathbb{Z}_6$ orbifold models with two Wilson lines, labelled models A1, A2 and B. We have obtained the complete spectra of massless states (plus KK excitations for these models in certain limits). As we now show, model A1 is the string equivalent to the orbifold GUT in sect. II B.

The following discussion relies greatly on the notation and discussion in appendices A and B. Briefly stated, the heterotic string combines a 10d superstring for right movers and a 26d bosonic string for left movers. However 16 of the 26 left-moving dimensions are compactified on the $\mathrm{E}_8\times\mathrm{E}_8$ root lattice. In order to obtain an effective 4d theory, we compactify six of the remaining ten dimensions on a symmetric orbifold defined by a six torus modded by a point group $\mathbb{Z}_6$ with the twist vector

$$\mathbf{v}_6 = \frac{1}{6}(1,2,-3), \tag{3.1}$$

i.e. the three compactified complex coordinates transform as $Z_i \to \exp(2\pi \mathrm{i}\mathbf{v}_6^i) Z_i$ under the twist. The embedding of orbifold twists in the gauge degrees of freedom is realized by gauge twists, $\mathbf{V}$, and lattice translations by discrete Wilson lines, $\mathbf{W}$. In abelian orbifolds these vectors simply shift the appropriate $\mathrm{E}_8\times\mathrm{E}_8$ roots.

To be definite, we choose the six torus as the Lie algebra root lattice $\mathrm{G}_2\oplus\mathrm{SU}_3\oplus\mathrm{SO}_4$, as shown in fig. 2. Denoting the basis of the lattice by $\mathbf{e} = \begin{pmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{pmatrix}\oplus\begin{pmatrix}\mathbf{e}_3\\ \mathbf{e}_4\end{pmatrix}\oplus\begin{pmatrix}\mathbf{e}_5\\ \mathbf{e}_6\end{pmatrix}$, whose inner product gives the Cartan matrix of the corresponding Lie algebra, the $\mathbb{Z}_6$ discrete symmetry can be realized

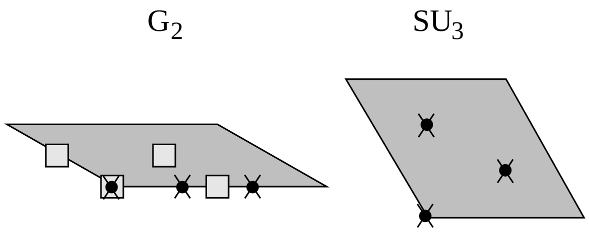

FIG. 2: Fundamental region of the root lattice $G_2 \oplus SU_3 \oplus SO_4$. The filled circles, crosses and squares represent fixed points in the T_1, $T_{2,4}$ and T_3 twisted sectors. See appendix B for further details.

by the Coxeter element,[5]

$$C = \begin{pmatrix} 2 & -1 \\ 3 & -1 \end{pmatrix} \oplus \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \oplus \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.2}$$

under which the basis is transformed to $C\mathbf{e}$. The Coxeter element has eigenvalues $\mathrm{e}^{\pm \mathrm{i}\pi/3}$, $\mathrm{e}^{\pm 2\mathrm{i}\pi/3}$ and $\mathrm{e}^{\pm \mathrm{i}\pi}$, thus the three two-dimensional sub-lattices have degree-6, 3, and 2 cyclic symmetries, and the corresponding numbers of fixed points are 1, 3 and 4.

There are three Kähler class moduli ($\mathcal{T}_{1,2,3}$), whose real parts parameterize the sizes of the three tori, and one complex structure modulus ($\mathcal{U}_3$), which parameterizes the shape of the third torus. Explicitly, $\mathrm{Re}\mathcal{T}_3 = 2RR' \sin\phi$, and $\mathcal{U}_3 = \frac{R}{R'}\,\mathrm{e}^{\mathrm{i}\phi}$, where R, R' are the lengths of the two axes of the SO_4-lattice and ϕ their relative angle. These moduli are arbitrary parameters. One may make the length of one axis (along which one puts the degree-2 Wilson line, $\mathbf{W}_2$), say R, large compared to the string length scale while keeping all other dimensions small. In this limit (for length scales larger than the string scale but smaller than the radius R), the low energy theory is effectively five dimensional.[6] The SO_4 lattice, on which only the $\mathbb{Z}_2$ sub-orbifold twist acts, has four fixed points. With only one degree-2 Wilson line, the fixed points split into two inequivalent classes, labelled by the winding number $n_2 = 0, 1$. Thus in our setup the fifth dimension is equivalent to the orbicircle $S^1/\mathbb{Z}_2$ where each of the two fixed points has a degree-2 degeneracy.

Note that we can reinterpret the $\mathbb{Z}_6$ models of appendix B in terms of the equivalent $\mathbb{Z}_2 \times \mathbb{Z}_3$ orbifold (where the $\mathbb{Z}_2$ ($\mathbb{Z}_3$) sub-orbifold twist acts on the G_2 and SO_4 (G_2 and SU_3) sub-lattices). This point of view is more useful for our comparisons with the orbifold GUTs in sect. II B. Labelling

[5] The Coxeter element is an inner automorphism of the lattice, composed of products of Weyl reflections of the corresponding root lattice. For example, the Coxeter element of G_2 is simply $s_1 s_2$ where s_1, s_2 are the two reflections with respect to planes orthogonal to the two simple roots. A generalized Coxeter element may also include outer automorphism of the lattice.

[6] It should be obvious that our construction can be generalized to 6d models, simply by taking both R and R' large compared to the string length scale. These models are related to 6d orbifold GUTs compactified on $T^2/\mathbb{Z}_2$.

a twisted sector in the $\mathbb{Z}_6$ model by T_k where $k = 1, 2, \cdots, 5$ and in the $\mathbb{Z}_2 \times \mathbb{Z}_3$ model by $T_{(k,l)}$ where $k = 0, 1$ and $l = 0, 1, 2$, then the correspondence between the twisted sectors in the $\mathbb{Z}_6$ and $\mathbb{Z}_2 \times \mathbb{Z}_3$ orbifolds is the following:

$\mathbb{Z}_6$ orbifold	T_1	T_2	T_3	T_4	T_5
$\mathbb{Z}_2 \times \mathbb{Z}_3$ orbifold	$T_{(1,2)}$	$T_{(0,1)}$	$T_{(1,0)}$	$T_{(0,2)}$	$T_{(1,1)}$

. (3.3)

The $T_{2,4}$ sectors, which will shortly be identified with the bulk states in the language of orbifold GUTs, have $k = 0,\ l = 1, 2$; therefore they are untwisted by the $\mathbb{Z}_2$ twist.

A. Model A1 from the $\mathbb{Z}_6$ orbifold compactification

We now examine model A1 of appendix B. Consider first the model with only the $\mathbb{Z}_3$ sub-orbifolding being imposed (i.e., with twist vector $\mathbf{v}_3 = 2\mathbf{v}_6$, gauge twist $\mathbf{V}_3 = 2\mathbf{V}_6$ and a degree-3 Wilson line $\mathbf{W}_3$, where $\mathbf{v}_6$, $\mathbf{V}_6$ and $\mathbf{W}_3$ are given in eqs. 3.1, B1 and B3), we find a 6d $\mathcal{N} = 2$ model with observable-sector gauge group E_6 (modulo abelian factors). Matter fields of the observable sector consist of 6d $\mathcal{N} = 2$ hypermultiplets in the following representations,

$$U\,\text{sectors}:\ \mathbf{27} + \overline{\mathbf{27}}, \qquad T\,\text{sectors}:\ 3 \times (\mathbf{27} + \overline{\mathbf{27}})\,. \tag{3.4}$$

The remaining $\mathbb{Z}_2$ twist acts as a space reversal on the third compactified complex dimension, $Z_3 \to -Z_3$. The $\mathbb{Z}_3$ models have two gravitini with the SO_8 momentum vectors, $\mathbf{r} = \frac{1}{2}(1, 1, 1, 1)$ and $\frac{1}{2}(1, -1, -1, 1)$, in the Ramond sector of the right-moving superstring (see appendix A for notation). Only one of them, $\mathbf{r} = \frac{1}{2}(1, 1, 1, 1)$, satisfies the $\mathbb{Z}_2$ projection, $\mathbf{r} \cdot \mathbf{v}_2 = \mathbb{Z}$. Hence the $\mathcal{N} = 2$ supersymmetry is broken to that of $\mathcal{N} = 1$ in 4d.

Gauge symmetry breaking induced by the $\mathbb{Z}_2$ orbifolding is as follows. The twist vector $\mathbf{v}_2$ is embedded in the gauge degrees of freedom in two different ways, with gauge twists $\mathbf{V}_2$ and $\mathbf{V}_2' = \mathbf{V}_2 + \mathbf{W}_2$ where $\mathbf{V}_2 = 3\mathbf{V}_6$ and $\mathbf{W}_2$ is given in eq. B3. E_6 generators in the Cartan-Weyl basis are transformed under the $\mathbb{Z}_2$ action as $E_{\mathbf{P}} \to \mathrm{e}^{2\pi \mathrm{i} \mathbf{P}\cdot\mathbf{V}_2} E_{\mathbf{P}}$ and $E_{\mathbf{P}} \to \mathrm{e}^{2\pi \mathrm{i} \mathbf{P}\cdot\mathbf{V}_2'} E_{\mathbf{P}}$, thus the linearly-realized gauge groups consist of roots satisfying $\mathbf{P} \cdot \mathbf{V}_2$ and $\mathbf{P} \cdot \mathbf{V}_2' = \mathbb{Z}$ respectively. The pattern of symmetry breaking in the observable sector can be summarized as follows:

$$\begin{tikzcd} & \mathrm{SO}_{10} \arrow[dr] & \\ \mathrm{E}_6 \arrow[ur] \arrow[dr] & & \mathrm{PS} \\ & \mathrm{SU}_6 \times \mathrm{SU}_{2\mathrm{R}} \arrow[ur] & \end{tikzcd} \tag{3.5}$$

At the final step we have the complete $\mathbb{Z}_6$ model with two discrete Wilson lines being imposed simultaneously; this gives the PS symmetry group in the 4d effective theory.

In these two inequivalent implementations of the $\mathbb{Z}_2$ twist the non-trivial matter fields of SO_{10} and $\mathrm{SU}_6 \times \mathrm{SU}_{2\mathrm{R}}$ are:

Sectors	SO_{10}	$\mathrm{SU}_6 \times \mathrm{SU}_{2\mathrm{R}}$
U_1	$\mathbf{16}$	$(\mathbf{15}, \mathbf{1})$
U_2	$\mathbf{10}$	$(\mathbf{6}, \mathbf{2})$
U_3	$\mathbf{16} + \overline{\mathbf{16}}$	$(\mathbf{20}, \mathbf{2})$
$T_{(0,1)}$	$2 \times \mathbf{16}_+ + \mathbf{10}_-$	$2(\overline{\mathbf{6}}, \mathbf{2})_+ + (\mathbf{15}, \mathbf{1})_-$
$T_{(0,2)}$	$\overline{\mathbf{16}}_- + 2 \times \mathbf{10}_+$	$(\mathbf{6}, \mathbf{2})_- + 2(\overline{\mathbf{15}}, \mathbf{1})_+$

(3.6)

where the subscripts $\pm$ represent *intrinsic parities*,

$$p = \gamma\phi\,. \tag{3.7}$$

p depends on the twist eigenvalue, γ, and the oscillator phase, ϕ; they are defined in appendix A. Note that $p = +$ for gauge and untwisted-sector states, and $p = +$ and $-$ have multiplicities 2 and 1 respectively for non-oscillator $T_{(01)}/T_{(02)}$ states.

Massless states in the untwisted and $T_{(0,1)}$, $T_{(0,2)}$ twisted sectors of model A1 are the intersections of those of the SO_{10} and $\mathrm{SU}_6 \times \mathrm{SU}_{2\mathrm{R}}$ models. This can be seen from the group branching rules. For example, the $T_{(0,1)}$-sector matter has the following branchings,

$$\begin{aligned} \mathrm{SO}_{10} &\rightarrow \mathrm{SU}_{4\mathrm{C}} \times \mathrm{SU}_{2\mathrm{L}} \times \mathrm{SU}_{2\mathrm{R}} \\ \mathbf{16}_+ &= (\mathbf{4}, \mathbf{2}, \mathbf{1})_+ + (\overline{\mathbf{4}}, \mathbf{1}, \mathbf{2})_+\,, \\ \mathbf{10}_- &= (\mathbf{6}, \mathbf{1}, \mathbf{1})_- + (\mathbf{1}, \mathbf{2}, \mathbf{2})_-\,, \end{aligned} \tag{3.8}$$

$$\begin{aligned} \mathrm{SU}_6 \times \mathrm{SU}_{2\mathrm{R}} &\rightarrow \mathrm{SU}_{4\mathrm{C}} \times \mathrm{SU}_{2\mathrm{L}} \times \mathrm{SU}_{2\mathrm{R}} \\ (\overline{\mathbf{6}}, \mathbf{2})_+ &= (\overline{\mathbf{4}}, \mathbf{1}, \mathbf{2})_+ + (\mathbf{1}, \mathbf{2}, \mathbf{2})_+\,, \\ (\mathbf{15}, \mathbf{1})_- &= (\mathbf{4}, \mathbf{2}, \mathbf{1})_- + (\mathbf{6}, \mathbf{1}, \mathbf{1})_- + (\mathbf{1}, \mathbf{1}, \mathbf{1})_-\,. \end{aligned} \tag{3.9}$$

The states in common, $2(\overline{\mathbf{4}}, \mathbf{1}, \mathbf{2})_+ + (\mathbf{6}, \mathbf{1}, \mathbf{1})_-$, agree with that of the T_2-twisted sector in eq. B6.

Massless fields in the other, i.e. $T_{(1,2)}(= T_1)$ and $T_{(1,0)}(= T_3)$, twisted sectors are the unions of those of the SO_{10} and $\mathrm{SU}_6 \times \mathrm{SU}_{2\mathrm{R}}$ models. Therefore there are two sets of states, furnishing complete representations of SO_{10} and $\mathrm{SU}_6 \times \mathrm{SU}_{2\mathrm{R}}$ respectively. For example, the T_1 sector of model A1 contains $(\mathbf{4}, \mathbf{2}, \mathbf{1}) + (\overline{\mathbf{4}}, \mathbf{1}, \mathbf{2})$ and $(\mathbf{4}, \mathbf{1}, \mathbf{1}) + (\mathbf{1}, \mathbf{2}, \mathbf{1})$, they are in the complete representations $\mathbf{16}$ of SO_{10} and $(\mathbf{6}, \mathbf{1})$ of $\mathrm{SU}_6 \times \mathrm{SU}_{2\mathrm{R}}$. In the notation of appendix A, these two sets of states have

quantum numbers $n_2 = 0$ and $n_2 = 1$. (These quantum numbers are the winding numbers along the direction where the $\mathbf{W}_2$ Wilson line is imposed.) The $n_2 = 0$ and $n_2 = 1$ fixed points are thus the SO_{10} and $SU_6 \times SU_{2R}$ branes in the orbifold GUT language.

B. Identifying orbifold parities in string theory

To a certain degree, the above E_6 heterotic model gives a string theoretical realization of the orbifold GUT in sect. II B. Better yet, we also achieve an understanding of the orbifold parities in terms of string theoretical quantities. Specifically, the analogue of orbifold parities, eq. 2.4, in our $\mathbb{Z}_6$ string models can be defined as follows [21]

$$P = p\mathrm{e}^{2\pi \mathrm{i}(\mathbf{P}\cdot\mathbf{V}_2 - \mathbf{r}\cdot\mathbf{v}_2)}\,, \qquad P' = p\mathrm{e}^{2\pi \mathrm{i}(\mathbf{P}\cdot\mathbf{V}'_2 - \mathbf{r}\cdot\mathbf{v}_2)}\,, \tag{3.10}$$

where $\mathbf{V}_2$ and $\mathbf{V}'_2$ are the two inequivalent gauge embeddings of the $\mathbb{Z}_2$ twist in sect. III A, and p is the intrinsic parity.

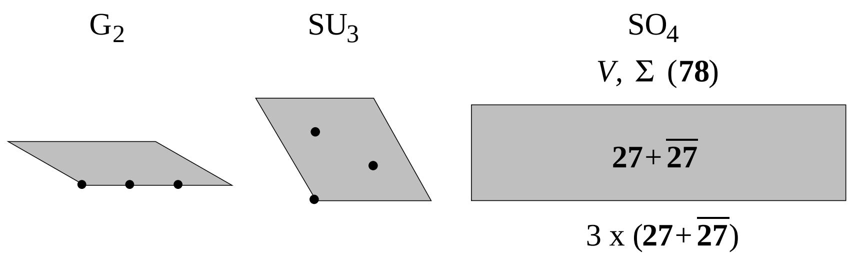

FIG. 3: $G_2 \oplus SU_3 \oplus SO_4$ lattice with $\mathbb{Z}_3$ fixed points. The fields V, Σ, and $\mathbf{27}(\in U_1) + \overline{\mathbf{27}}(\in U_2)$ are bulk states from the untwisted sectors. On the other hand, $3 \times (\mathbf{27} + \overline{\mathbf{27}})$ are "bulk" states located on the $T_{(0,1)}/T_{(0,2)}$ twisted sector (G_2, SU_3) fixed points.

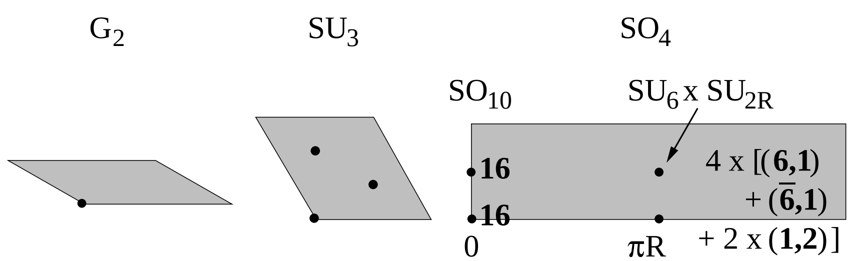

FIG. 4: $G_2 \oplus SU_3 \oplus SO_4$ lattice with $\mathbb{Z}_6$ fixed points. The $T_{(1,1)}/T_{(1,2)}$ twisted sector states sit at these fixed points.

These parities can be deduced from the generalized Gliozzi-Scherk-Olive (GSO) projector [10, 28], as in the paragraphs after eq. A14. Since the terms in the exponents, $\mathbf{P} \cdot \mathbf{V}_2 - \mathbf{r} \cdot \mathbf{v}_2$ and

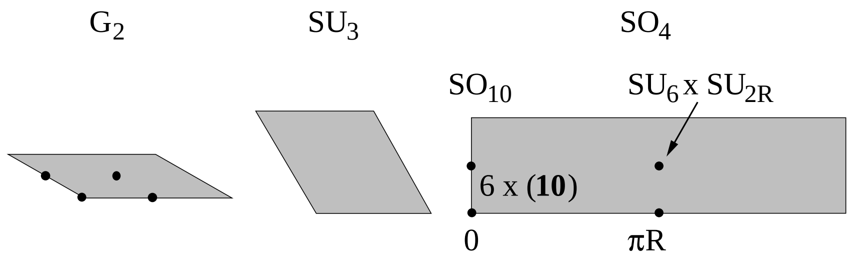

FIG. 5: $G_2 \oplus SU_3 \oplus SO_4$ lattice with $\mathbb{Z}_2$ fixed points. The $T_{(1,0)}$ twisted sector states sit at these fixed points.

$\mathbf{P}\cdot\mathbf{V}_2' - \mathbf{r}\cdot\mathbf{v}_2$, take integral or half-integral values, P and P' are either $+$ or $-$. The orbifold translation corresponds to the difference in P and P', i.e. $T = \mathrm{e}^{2\pi \mathrm{i}\mathbf{P}\cdot\mathbf{W}_2}$. The P, P' and T in string models have exactly the same properties as that of the orbifold GUTs.

Evidently, in the E_6 orbifold GUT model of sect. II B states supported at the SO_{10} and $SU_6 \times SU_{2R}$ branes are those with parities $P = +$ and $P' = +$, and states in the 4d effective theory are those with parities $P = P' = +$; this agrees with the string theoretical interpretation, since the parities in eq. 3.10 are nothing but the required GSO projections for the gauge, untwisted and $T_{(01)}/T_{(02)}$ sector states (i.e. the bulk states) in string models. (The massless states, i.e. $P = P' = +$ modes from bulk and $T_{(0,1)}/T_{(0,2)}$ twisted sectors are shown in fig. 3.) From information gathered in sect. III A and appendix B, we can also deduce the P and P' parities for the various bulk matter states. They are listed in table I.

KK masses for these bulk states can also be derived in string models. The mode expansions of the coordinates corresponding to the SO_4 lattice are $X^i_{L,R} = x^i_{L,R} + p^i_{L,R}(\tau \pm \sigma) + \text{oscillator terms}$, with p^i_L, p^i_R given by eq. A3. The $\mathbb{Z}_2$ action maps m to $-m$, n to $-n$ and $\mathbf{W}_2$ to $-\mathbf{W}_2$, so physical states must contain linear combinations, $|m,n\rangle \pm |-m,-n\rangle$; the eigenvalues ± 1 correspond to the first $\mathbb{Z}_2$ parity of the orbifold GUT models. The second embedding corresponds to a non-trivial Wilson line; it shifts the KK level by $m \to m + \mathbf{P}\cdot\mathbf{W}_2$. Since $2\mathbf{W}_2$ is a vector of the integral $E_8 \times E_8$ lattice, the shift $\mathbf{P}\cdot\mathbf{W}_2$ must be an integer or half-integer. In the orbifold GUT limit when the winding modes and the KK modes in the short direction of SO_4 decouple, eq. A3 reproduces the field theoretical mass formula in eq. 2.2.

As seen in sect. III A, matter states in the $T_{(1,1)}/T_{(1,2)}$ and $T_{(1,0)}$ twisted sectors, which may be identified with the first two families, are localized on the two inequivalent fixed points in the SO_4 lattice. They are the SO_{10} and $SU_6 \times SU_{2R}$ brane states (See figs. 4 and 5). These twisted-sector states are more tightly constrained than their orbifold GUT counterparts. In orbifold GUT models the only consistency requirement is the chiral anomaly cancellation, thus one can add

TABLE I: Parities for the bulk states in model A1, computed from eq. 3.10. The states have been decomposed to the PS irreducible representations.

Multiplicities	States	P	P'	States	P	P'
1	$V(\mathbf{15},\mathbf{1},\mathbf{1})$	+	+	$\Sigma(\mathbf{15},\mathbf{1},\mathbf{1})$	−	−
1	$V(\mathbf{1},\mathbf{3},\mathbf{1})$	+	+	$\Sigma(\mathbf{1},\mathbf{3},\mathbf{1})$	−	−
1	$V(\mathbf{1},\mathbf{1},\mathbf{3})$	+	+	$\Sigma(\mathbf{1},\mathbf{1},\mathbf{3})$	−	−
1	$V(\mathbf{6},\mathbf{2},\mathbf{2})$	+	−	$\Sigma(\mathbf{6},\mathbf{2},\mathbf{2})$	−	+
1	$V(\mathbf{4},\mathbf{2},\mathbf{1})$	−	+	$\Sigma(\mathbf{4},\mathbf{2},\mathbf{1})$	+	−
1	$V(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})$	−	−	$\Sigma(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})$	+	+
1	$V(\overline{\mathbf{4}},\mathbf{2},\mathbf{1})$	−	+	$\Sigma(\overline{\mathbf{4}},\mathbf{2},\mathbf{1})$	+	−
1	$V(\mathbf{4},\mathbf{1},\mathbf{2})$	−	−	$\Sigma(\mathbf{4},\mathbf{1},\mathbf{2})$	+	+
1	$H(\mathbf{4},\mathbf{2},\mathbf{1})$	+	+	$H^c(\overline{\mathbf{4}},\mathbf{2},\mathbf{1})$	−	−
1	$H(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})$	+	−	$H^c(\mathbf{4},\mathbf{1},\mathbf{2})$	−	+
1	$H(\mathbf{6},\mathbf{1},\mathbf{1})$	−	+	$H^c(\mathbf{6},\mathbf{1},\mathbf{1})$	+	−
1	$H(\mathbf{1},\mathbf{2},\mathbf{2})$	−	−	$H^c(\mathbf{1},\mathbf{2},\mathbf{2})$	+	+
2	$H(\mathbf{4},\mathbf{2},\mathbf{1})_+$	+	−	$H^c(\overline{\mathbf{4}},\mathbf{2},\mathbf{1})_+$	−	+
2	$H(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})_+$	+	+	$H^c(\mathbf{4},\mathbf{1},\mathbf{2})_+$	−	−
2	$H(\mathbf{6},\mathbf{1},\mathbf{1})_+$	−	−	$H^c(\mathbf{6},\mathbf{1},\mathbf{1})_+$	+	+
2	$H(\mathbf{1},\mathbf{2},\mathbf{2})_+$	−	+	$H^c(\mathbf{1},\mathbf{2},\mathbf{2})_+$	+	−
1	$H(\mathbf{4},\mathbf{2},\mathbf{1})_-$	−	+	$H^c(\overline{\mathbf{4}},\mathbf{2},\mathbf{1})_-$	+	−
1	$H(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})_-$	−	−	$H^c(\mathbf{4},\mathbf{1},\mathbf{2})_-$	+	+
1	$H(\mathbf{6},\mathbf{1},\mathbf{1})_-$	+	+	$H^c(\mathbf{6},\mathbf{1},\mathbf{1})_-$	−	−
1	$H(\mathbf{1},\mathbf{2},\mathbf{2})_-$	+	−	$H^c(\mathbf{1},\mathbf{2},\mathbf{2})_-$	−	+

arbitrary numbers of vector-like representations to the branes. String models have to satisfy more stringent modular invariance conditions [4, 26] (of course, one-loop modular invariance guarantees the model is anomaly free, up to a possible anomalous abelian factor [29]), which also constrains any additional matter in vector-like representations.

C. Other models

In this subsection, we discuss the other two models in appendix B in the orbifold GUT language, for completeness. These models do not have matter-Higgs couplings at the renormalizable level, and may have limited phenomenological interest.

Model A2 has already been analyzed in ref. [21]. In the 5d bulk, it has an $\mathrm{SO}_{10} \times \mathrm{SU}_2$ gauge

symmetry and the following set of matter states,

$$U\,\text{sectors}:(\mathbf{16},\mathbf{1})+(\mathbf{1},\mathbf{2}),\qquad T\,\text{sectors}:3(\mathbf{16},\mathbf{1})+6(\mathbf{10},\mathbf{1})+15(\mathbf{1},\mathbf{2})\,. \tag{3.11}$$

The bulk gauge group is unbroken at the fixed point $x^5=0$ and broken to the PS group at $x^5=\pi R$ respectively, the states supported at these two points are

Sectors	$SO_{10}\times SU_2$	PS
U_1	$(\mathbf{16},\mathbf{1})$	$(\mathbf{4},\mathbf{2},\mathbf{1})$
U_2	$(\mathbf{1},\mathbf{2})$	$(\mathbf{4},\mathbf{1},\mathbf{2})$
U_3	$-$	$(\mathbf{6},\mathbf{2},\mathbf{2})$
$T_{(0,1)}$	$2(\mathbf{16},\mathbf{1})_++2(\mathbf{10},\mathbf{1})_-$	$(\mathbf{6},\mathbf{1},\mathbf{1})_-+2(\mathbf{1},\mathbf{2},\mathbf{2})_+$
	$+2(\mathbf{1},\mathbf{2})_++4(\mathbf{1},\mathbf{2})_-$	$+2(\mathbf{6},\mathbf{1},\mathbf{1})_++(\mathbf{1},\mathbf{2},\mathbf{2})_-$
		$+(\mathbf{4},\mathbf{2},\mathbf{1})_-+2(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})_+$
$T_{(0,2)}$	$(\overline{\mathbf{16}},\mathbf{1})_-+4(\mathbf{10},\mathbf{1})_+$	$2(\mathbf{6},\mathbf{1},\mathbf{1})_++(\mathbf{1},\mathbf{2},\mathbf{2})_-$
	$+8(\mathbf{1},\mathbf{2})_++(\mathbf{1},\mathbf{2})_-$	$+(\mathbf{6},\mathbf{1},\mathbf{1})_-+2(\mathbf{1},\mathbf{2},\mathbf{2})_+$
		$+(\mathbf{4},\mathbf{1},\mathbf{2})_-+2(\overline{\mathbf{4}},\mathbf{2},\mathbf{1})_+$

. (3.12)

Model B is similar to model A1, with an E_6 bulk group and the same set of bulk states as in eq. 3.4. The E_6 group is broken to SO_{10} and $SU_6\times SU_{2L}$ respectively at the two fixed points, and the matter states are

Sectors	SO_{10}	$SU_6\times SU_{2L}$
U_1	$\overline{\mathbf{16}}$	$(\mathbf{6},\mathbf{2})$
U_2	$\mathbf{10}$	$(\mathbf{15},\mathbf{1})$
U_3	$\mathbf{16}+\overline{\mathbf{16}}$	$(\mathbf{20},\mathbf{2})$
$T_{(0,1)}$	$2\times\overline{\mathbf{16}}_++\mathbf{10}_-$	$(\mathbf{6},\mathbf{2})_-+2(\overline{\mathbf{15}},\mathbf{1})_+$
$T_{(0,2)}$	$\mathbf{16}_-+2\times\mathbf{10}_+$	$2(\overline{\mathbf{6}},\mathbf{2})_++(\mathbf{15},\mathbf{1})_-$

. (3.13)

The 4d effective theories of models A2 and B have a PS symmetry, and the complete matter content is listed in appendix B. Similar to model A1, matter fields in the untwisted and T_2/T_4 twisted sectors can be traced back to the states in the above two tables, by using appropriate group branching rules. We also note that both models contain two families of chiral matter from the fixed point at $x^5=0$ and one family from the bulk, a common feature to all our models. This feature predicts a non-abelian dihedral D_4 family symmetry – *a novelty in string model building* – as we will see in the next section.

D. The color-triplet problem

A major motivation for constructing orbifold GUT models is the well-known doublet-triplet splitting problem in conventional 4d GUTs. Orbifold GUTs solve this problem by assigning appropriate orbifold parities to the Higgs doublets and triplets, such that the triplets are automatically projected out of the effective theory [19].[7] We have already seen this in the SO_{10} model in sect. II A. This mechanism, however, usually cannot be trivially implemented in heterotic models.

The difficulty is largely due to the intricate nature of string models. These models need to satisfy delicate modular invariance consistency conditions [4, 26] and are physically more constrained than the orbifold GUTs. Before imposing any Wilson line, the $\mathbb{Z}_6$ models of eqs. B1 and B2 always contain the $\mathbf{10}$ representation of SO_{10} simultaneously in several sectors. We find it is impossible to design modular-invariant Wilson lines to fulfill the following requirements: (a) break the gauge group to PS in 4d, (b) give rise to three chiral families, and (c) eliminate the color triplets altogether. Furthermore, the $\mathbf{10}$ representation of the T_3 sector in model A1 does not suffer from additional projections even when the $\mathbf{W}_2$ Wilson line is turned on. Indeed, it simply decomposes to the $(\mathbf{6},\mathbf{1},\mathbf{1})+(\mathbf{1},\mathbf{2},\mathbf{2})$ representations under the PS group.

Although the presence of many color triplets is a nuisance, one $\mathbf{3}+\overline{\mathbf{3}}$ pair may be necessary to facilitate the breaking of PS to the SM gauge group, as illustrated in the E_6 model in sect. II B. Moreover, it is not entirely clear that the color triplets in our models pose the same problems as in conventional GUTs. Indeed, although there are color triplets (those of the T_1/T_3 twisted sectors) with doublet companions having exactly the same quantum numbers, in general we also have $(\mathbf{6},\mathbf{1},\mathbf{1})$ and $(\mathbf{1},\mathbf{2},\mathbf{2})$ states with different quantum numbers, in all three models. The usual doublet-triplet problem does not necessarily apply for the second situation. We need to check whether it is possible to make all color triplets sufficiently heavy while (hopefully) keeping one MSSM Higgs-doublet pair light; this requires a better understanding of the effective actions of our models and will be examined for model A1 in the next section.

IV. PHENOMENOLOGY OF MODEL A1

We have seen in the previous sections that the E_6 orbifold GUT and heterotic A1 model match nicely in the low energy regime. In this section, we study some phenomenological issues for the A1 model. We first study gauge coupling unification, relying on a simplification due to the correspon-

[7] Of course, more conventional field theoretical mechanisms [30] have been widely studied in the literature.

dence between the field and string theoretical models. We then study Yukawa couplings (including both renormalizable and non-renormalizable couplings), concentrating on several immediate phenomenological questions: (i) breaking of the PS symmetry to that of the SM, (ii) mass generation for the color triplets, (iii) proton stability, and (iv) matter-Higgs Yukawa couplings. These couplings are introduced by hand in the orbifold GUTs. In string models we no longer enjoy the same kind of freedom. In fact, these couplings are determined by string selection rules (reviewed in appendix C 1). The low energy phenomenology also depends crucially on the flat directions of the effective $\mathcal{N} = 1$ model. However, we do not attempt to solve this complicated problem here.

A. Gauge coupling unification and proton decay

As discussed earlier, since the first two families are located on the SO_{10} brane, proton decay constraints require that the 5d compactification scale M_c be greater than $O(10^{16})$ GeV. However all these GUT scale thresholds must be consistent with low energy gauge coupling unification. Consider the solution to the 5d renormalization group (RG) equations, i.e., the GQW equations, in the orbifold GUT limit,[8] given by

$$\begin{aligned}\frac{2\pi}{\alpha_i(\mu)} \simeq\ & \frac{2\pi}{\alpha_{\text{string}}} + b_i^{\text{MSSM}} \log\frac{M_{\text{PS}}}{\mu} + (b_{++}^{\text{PS}} + b_{\text{brane}})_i \log\frac{M_{\text{string}}}{M_{\text{PS}}} \\ & - \frac{1}{2}(b_{++}^{\text{PS}} + b_{--}^{\text{PS}})_i \log\frac{M_{\text{string}}}{M_c} + b^{\text{E}_6}\left(\frac{M_{\text{string}}}{M_c} - 1\right),\end{aligned} \tag{4.1}$$

where M_{PS} is the PS breaking scale and α_{string} is the gauge coupling at the string scale. In addition, in the weakly coupled heterotic string we have the boundary condition

$$\frac{2\pi}{\alpha_{\text{string}}} = \frac{\pi}{4}\left(\frac{M_{\text{Pl}}}{M_{\text{string}}}\right)^2 + \frac{1}{2}\,\Delta^{\text{univ}}, \tag{4.2}$$

where the first term is the tree level result and the second is a universal one loop stringy correction. The latter correction depends on the value of the $\mathcal{T}_3$, $\mathcal{U}_3$ moduli. Following ref. [36] we see that Δ^{univ} is a finite function of its argument (with a mild singularity when $\mathcal{T}_3 = \mathcal{U}_3$, modulo $\text{PSL}(2,\mathbb{Z})$ transformations). Since the universal correction is not significant, we use the tree level formula in the following.

[8] In principle these equations can be derived from a string theory calculation, following refs. [31, 32, 33, 34, 35]. However, it is difficult to obtain the GQW equations in analytic form in string models with discrete Wilson lines [35], which makes the calculation less practical for our purposes. Instead, in deriving eq. 4.1, we have worked in the orbifold GUT limit, and assumed the most important contributions to the gauge threshold corrections come from the KK tower of the large dimension of the SO_4 lattice, with a physical cutoff at the string scale, M_{string}. See appendix D for more details.

Eq. 4.1 can be compared mathematically to the 4d equations given by

$$\frac{2\pi}{\alpha_i(\mu)} \simeq \frac{2\pi}{\alpha_{\rm GUT}} + b_i^{\rm MSSM} \log \frac{M_{\rm GUT}}{\mu} + 6\,\delta_{i3}, \tag{4.3}$$

where $M_{\rm GUT} \simeq 3 \times 10^{16}$ GeV, $\alpha_{\rm GUT}^{-1} \simeq 24$ and we have included a threshold correction at $M_{\rm GUT}$, required in order to fit the low energy data.

With the bulk field parities given in table I, we find the beta function coefficients, $\frac{1}{2}(b_{++}^{\rm PS} + b_{--}^{\rm PS}) = (1,1,1)$, $\quad b_{++}^{\rm PS} = (\frac{9}{5}, -3, -3)$. The brane contributions include that of the two PS families, $b_{\rm brane} = (4,4,4)$, and those from extra matter fields, $n'_{\mathbf{6}+\overline{\mathbf{6}}}(\frac{2}{5},1,1) + n_{\mathbf{10}}(1,1,1) + n'_{\mathbf{2}_R}(\frac{3}{10},0,0)$ (with $n'_{\mathbf{6}+\overline{\mathbf{6}}} \leq 4$, $n_{\mathbf{10}} \leq 6$ and $n'_{\mathbf{2}_R} \leq 12$). Equating the difference $2\pi/\alpha_3(\mu) - 2\pi/\alpha_2(\mu)$, eqs. 4.1 and 4.3 gives

$$M_{\rm PS} \simeq {\rm e}^{-3/2} M_{\rm GUT} \simeq 7 \times 10^{15} \text{ GeV}. \tag{4.4}$$

Then equating the difference $2\pi/\alpha_2(\mu) - 2\pi/\alpha_1(\mu)$, we find

$$\log \frac{M_{\rm string}}{M_{\rm GUT}} \simeq \frac{8 - 3n'_{\mathbf{2}_R} + 6n'_{\mathbf{6}+\overline{\mathbf{6}}}}{32 + 2n'_{\mathbf{2}_R} - 4n'_{\mathbf{6}+\overline{\mathbf{6}}}}, \tag{4.5}$$

which results in a maximum value for $M_{\rm string}$ for $n'_{\mathbf{2}_R} = 0$, $n'_{\mathbf{6}+\overline{\mathbf{6}}} = 4$, given by

$$M_{\rm string}^{\rm max} \simeq {\rm e}^2 M_{\rm GUT} \simeq 2 \times 10^{17} \text{ GeV}. \tag{4.6}$$

Finally we have

$$\frac{\alpha_{\rm GUT}}{\alpha_{\rm string}} - 1 \simeq \frac{\alpha_{\rm GUT}}{2\pi} \left[\log \frac{M_{\rm GUT}}{M_c} - (n_{10} + n'_{\mathbf{6}+\overline{\mathbf{6}}}) \left(\log \frac{M_{\rm string}}{M_{\rm GUT}} + \frac{3}{2} \right) \right]. \tag{4.7}$$

Using $M_{\rm string} = M_{\rm string}^{\rm max}$ and eq. 4.2 for the tree level heterotic string boundary condition, we find there is *no* solution, consistent with $M_{\rm string} > M_c \simeq M_{\rm GUT} > M_{\rm PS}$. The problem is that the value of $\alpha_{\rm string}$ given by eq. 4.2 is much too small ($\alpha_{\rm string} \ll \alpha_{\rm GUT}$) and it cannot be obtained by logarithmic running above the compactification scale (note, $b^{\rm E_6} = 0$ and thus there is no power-law running). This problem suggests that non-trivial (perhaps non-perturbative) string boundary conditions are required for consistency.

We have considered the 11d Hořava-Witten extension [37] of the perturbative heterotic string boundary condition given by [38]

$$\frac{2\pi}{\alpha_{\rm string}} = \frac{1}{2(4\pi)^{5/3} M_* \rho} \left(\frac{M_{\rm Pl}}{M_*} \right)^2, \tag{4.8}$$

where M_* is given in terms of the 11d Newton's constant by $\kappa^{2/3} = M_*^{-3}$ and ρ is the size of the eleventh dimension. Now using eq. 4.8, we find solutions for $M_{\rm string} \simeq M_* = 2M_{\rm GUT}$, $M_c \simeq M_{\rm PS} \simeq$

$e^{-3/2}M_{\rm GUT}$ with $n'_{\mathbf{2}_R} = n'_{\mathbf{6}+\overline{\mathbf{6}}} = 4$ and $M_*\rho \simeq 2$. Of course, this solution provides an enhanced proton decay rate due to dimension-6 operators with the dominant decay mode $p \to e^+\pi^0$. The decay rate for dimension-6 operators is given by [39]

$$\begin{aligned}\tau(p \to e^+\pi^0) &\simeq 1.25 \times 10^{36} \left(\frac{M_X}{3 \times 10^{16}\ \text{GeV}}\right)^4 \left(\frac{0.015\ \text{GeV}^3}{\beta_{\text{lattice}}}\right)^2\ \text{yrs} \\ &\simeq 3 \times 10^{33} \left(\frac{0.015\ \text{GeV}^3}{\beta_{\text{lattice}}}\right)^2\ \text{yrs},\end{aligned} \tag{4.9}$$

where β_{lattice} is an input from lattice calculations of the three quark matrix element.[9] Recent results give a range of central values $\beta_{\text{lattice}} = 0.007 - 0.015$ [40]. Note, the present experimental bound for this decay mode from Super-Kamiokande is 5.7×10^{33} years at 90% confidence levels [41]. Thus this prediction is not yet excluded by the data, but it should be observed soon.

B. Yukawa couplings

1. *PS symmetry breaking, mass generation for color-triplets and proton stability*

To successfully break the PS symmetry to that of the SM and generate mass for unwanted color triplet states, the 5d E_6 heterotic model should contain non-trivial couplings of the form in eq. 2.12. The model, however, contains additional color triplets. They could, in principle, develop mass through non-trivial Yukawa couplings to, say, singlet fields. In order to verify if this is a possibility, we need to know whether the required couplings exist in the 4d effective theory of the string model. For this purpose we are particularly interested in non-trivial couplings containing PS invariant operators, $(\mathbf{6},\mathbf{1},\mathbf{1})(\mathbf{6},\mathbf{1},\mathbf{1})$, $(\mathbf{6},\mathbf{1},\mathbf{1})(\mathbf{4},\mathbf{1},\mathbf{2})(\mathbf{4},\mathbf{1},\mathbf{2})$ and $(\mathbf{6},\mathbf{1},\mathbf{1})(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})$, that are allowed by string selection rules. (In field theory, all gauge invariant operators would be allowed.) These rules are reviewed in appendix C 1 and the relevant operators are given in eqs. C13 – C16.

Cubic, renormalizable, couplings in model A1 are determined in eq. C11, they contain the following operators of interest (we label the fields according to table II),

$$\mathcal{W}^{(3)} \supset S_1(C_3)_{A\alpha}(C_3)_{B\beta} + C_1\overline{\chi}^c_1\overline{\chi}^c_2 + (C_3)_{A\alpha}f^c_B\chi^c_\beta + (C_4)_\alpha\chi^c_\beta f^c_3 + (C_4)_\alpha f^c_A f^c_B, \tag{4.10}$$

where $\alpha, \beta = 1, 2$ labels the two $\gamma = 1$ eigenstates in the $T_{2,4}$ sectors, $A, B = 1, 2$ indicate degeneracies associated with the n'_2 winding number (which corresponds to a hidden S_2 permutation symmetry). The string selection rules require $\alpha + \beta = 0 \bmod 2$, $A + B = 0 \bmod 2$, and for the last

---

[9] To obtain this result we have taken the decay rate for SU_5 [39] and multiplied the amplitude by an additional factor of two to account for the extra gauge exchange present in SO_{10}.

term α is unrestricted. Apparently these couplings are not sufficient to break the PS symmetry and give mass to all the color triplets contained in the $(\mathbf{6},\mathbf{1},\mathbf{1})$ and $(\mathbf{4},\mathbf{1},\mathbf{2})+(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})$ states. Thus, to achieve our goal, we must also take into consideration higher-dimensional operators. Some of them are listed in appendix C 2. (Of course, there are many more operators with even higher dimensions. It is not obvious that "stringy zeroes" exist. We omitted these operators. They may or may not disrupt the following discussion.)

A few observations are appropriate. The fields C_i, $i = 1, \cdots, 4$ transform as a $\mathbf{3}_{-2/3} + \overline{\mathbf{3}}_{2/3}$ under $\mathrm{SU}_{3\mathrm{C}} \times \mathrm{U}_{1\mathrm{Y}}$ or as $\overline{D}^c + \overline{D}$ where $\overline{D}$ has the quantum number of an anti-down quark. These color triplets have multiplicities [in brackets], $C_1[1],\ C_2[2],\ C_3[4],\ C_4[2]$. In addition $C_2,\ C_3$ appear in complete SO_{10} **10**-plets. We also have states contained in $\chi^c_\alpha,\ \overline{\chi}^c_{1,2}$ with quantum numbers of $\overline{D},\ \overline{D}^c,\ \overline{U},\ \overline{U}^c,\ \overline{E}, \overline{E}^c$. Finally we have the color triplet states in $q_1,\ q_2,\ \overline{q}_1,\ \overline{q}_2$ with multiplicity 2 each. These latter are exotic states with fractional charge $\pm 1/2$ for the color singlets and $\pm 1/6$ for the color triplets.

For the exotic states we find operators of the form $\overline{q}_1(q_1 + q_2)$ multiplied by products of singlets S^n (up to sixth order), but no operators of the form $\overline{q}_2(q_1 + q_2)S^n$ to order $n = 9$. Hence $\overline{q}_2$ and one linear combination of $q_1,\ q_2$ remain massless. This is a serious problem for the model, since these states are absolutely stable and should have been observed. It remains to be seen whether the operator $\overline{q}_2(q_1 + q_2)S^n$ is generated at order $n \geq 10$ or if it is forbidden by the string selection rules to all orders.

Now consider the fields C_i, $i = 1, \cdots, 4$ and $\chi^c_\alpha,\ \overline{\chi}^c_{1,2}$. In this sector we need to both find a way of spontaneously breaking PS to the SM, as well as giving all color triplets mass. A related issue is the potential problem of rapid proton decay mediated by these color triplets. In particular we must eliminate or greatly suppress the following baryon/lepton-number violating effective operators

$$f f f^c \langle \chi^c S^n \rangle \Longrightarrow QL\overline{D} + LL\overline{E}, \qquad f^c f^c f^c \langle \chi^c S^n \rangle \Longrightarrow \overline{U}\,\overline{D}\,\overline{D}. \tag{4.11}$$

Note we have checked that these operators are not generated prior to integrating out the color triplets, for $n \leq 3$.[10] Nevertheless there is a danger that they will be generated in the effective theory below the color triplet mass. In fact, consider the renormalizable couplings in eq. 4.10. It is evident that an effective mass term of the form $\langle S \rangle (C_4)^2$ combined with the coupling $C_4(\chi^c_\alpha f^c_3 +$

[10] Of course, it would be better check to any order in n, or better yet, find a symmetry which forbids them to all orders.

$f^c_A f^c_B$) leads to the effective operator of the form

$$\frac{1}{\langle S\rangle}\chi^c_\alpha f^c_3 f^c_A f^c_B = \frac{M_{\rm PS}}{\langle S\rangle}\overline{U}_3\overline{D}_A\overline{D}_B. \tag{4.12}$$

Similarly, additional baryon/lepton-number violating operators are obtained from an effective C_3C_4 mass term. These baryon-number violating operator may be phenomenologically acceptable *if* the coefficient is sufficiently small. However it seems prudent to eliminate the offending mass terms by choosing a vacuum configuration where the appropriate scalar vevs vanish. For example, given the superpotential terms in appendix C, eq. C13, we demand that the following vevs (i.e. the coefficients of $(C_4)^2$ and C_3C_4) vanish $S_2S_{24} + \mathcal{S}^{(2)}_1\mathcal{S}^{(2)}_2 = \mathcal{S}^{(2)}_1 + S_2S_9S_{24} + S_1\mathcal{S}^{(2)}_2 + S_{19}S_{25}\mathcal{S}^{(2)}_2 = 0$.

In the following we consider the possibility of obtaining a baryon/lepton-number conserving low energy effective theory.[11] As a possible proof of existence, we suggest an ansatz where the only singlets with non-vanishing vevs are

$$\{S_1\}, \quad \{S_{10}\mathcal{S}^{(2)}_4\}, \quad [S_6S_7S_{14}S_{18}], \quad [S_7S_{14}S_{16}(S_6S_{26} + S_2S_{27})], \quad \{S_{26}\}, \tag{4.13}$$

where the curly braces represent classes of singlets with the same transformation properties under all string symmetries and we use square brackets when we explicitly present a finite set of fields in the same class. The corresponding superpotential from appendix C 2 is

$$\begin{aligned}\mathcal{W} \supset \quad & \Big(\{S_1\}(C_3)^2 + \{S_2S_{22}\mathcal{S}^{(2)}_9\}C_2C_3 + \{S_{10}\mathcal{S}^{(2)}_4\}C_1C_4\Big)\\
&+\Big(C_1\overline{\chi}^c_1\overline{\chi}^c_2 + C_3 f^c_A\chi^c_\alpha + C_4(\chi^c_\alpha f^c_3 + f^c_A f^c_B)\Big)\\
&+\Big(\{S_2S_{24}S_{26}\}C_4(\overline{\chi}^c_1)^2 + [S_{10}\mathcal{S}^{(2)}_1 + S_6S_7S_{14}S_{18} + S_1S_{10}\mathcal{S}^{(2)}_2 + S_2\mathcal{S}^{(3)}_4]C_4(\overline{\chi}^c_2)^2\\
&\qquad +[S_2S_3S_{12}\mathcal{S}^{(2)}_3 + S_7S_{14}S_{16}(S_6S_{26} + S_2S_{27})]C_4\chi^c_\alpha\chi^c_\beta\Big) + \{S_{26}\}\chi^c_\alpha\overline{\chi}^c_1.\end{aligned} \tag{4.14}$$

Note the coefficient of the C_2C_3 term and the first term linear in C_4 vanishes due to the vevs we have chosen, but there may be other higher-dimensional terms which replace them. For example, the first element in the second term linear in C_4 also vanishes, but the other terms may be non-zero.

The following vevs are assumed to vanish.

$$[S_2S_{24} + \mathcal{S}^{(2)}_1\mathcal{S}^{(2)}_2], \qquad\qquad [\mathcal{S}^{(2)}_1 + S_2S_9S_{24} + S_1\mathcal{S}^{(2)}_2 + S_{19}S_{25}\mathcal{S}^{(2)}_2],$$

[11] The conventional wisdom of field theory is to use an R-parity (or family reflection symmetry) to eliminate these baryon/lepton-number violating operators. (The R-parity has a bonus of predicting a generic stable neutral fermionic superpartner, which makes it even more appealing phenomenologically.) Although from the start our string models contain several discrete R symmetries at the level of 4d effective action, it is not clear a priori whether any of them can survive symmetry breaking. Note that in our models an unbroken R parity (which is capable of distinguishing the Higgs fields, χ^c, from the matter, f^c) does not exist because both $C_4\chi^c_\alpha f^c_3$ and $C_4 f^c_A f^c_B$ couplings are allowed by string selection rules at the renormalizable level.

$$
\begin{aligned}
&\{S_2S_9S_{22}\}, && [S_{10}\mathbb{S}_3^{(2)} + S_{13}\mathbb{S}_2^{(2)} + S_1S_{10}S_{12}S_{32} + S_9S_{10}\mathbb{S}_4^{(2)}],\\
\{S_9\} \equiv S_9 + S_1S_{10}S_{21}S_{22} + &\mathbb{S}_2^{(2)}S_{11}^2 + \cdots, && [\mathbb{S}_9^{(2)} + S_{10}(S_{30}\mathbb{S}_5^{(2)} + S_{13}\mathbb{S}_4^{(2)})],\\
&[\mathbb{S}_2^{(2)} + S_{26}\mathbb{S}_2^{(3)}], && [S_9\mathbb{S}_9^{(2)} + S_{13}(S_{10}\mathbb{S}_3^{(2)} + S_{13}\mathbb{S}_2^{(2)})],\\
&\{S_{10}S_{13}S_{21}S_{22}\}, && \{S_9S_{10}S_{12}S_{32}\}, \qquad \{S_{10}\mathbb{S}_1^{(3)}\}
\end{aligned} \tag{4.15}
$$

With this choice we guarantee, at least to the order we have checked, that we do not generate baryon/lepton-number violating operators in the low energy theory obtained by integrating out the color triplets. A self-consistent solution to the necessary set of vevs is given by

$$
S_4 = S_9 = S_{11} = S_{13} = S_{17} = S_{21} = S_{24} = S_{25} = S_{29} = S_{30} = S_{32} = 0, \tag{4.16}
$$

and all other vevs non-zero. In addition we require $[S_5S_{33} + S_{10}S_{26}] = 0$, which may or may not require fine-tuning. Unfortunately we are not able to identify a symmetry which would extend this result to all orders in string perturbation. This is a serious problem for the model.

The first term in parentheses of eq. 4.14 gives mass to the color triplets C_2, C_3, C_1 and one triplet in C_4. Since the doublets h_2 and h_3 have exactly the same quantum numbers as that of C_2 and C_3, they acquire the same mass and also decouple from the low energy effective theory. In this way, we obtain just one doublet field h_1 (from the U_2 sector). It is a good candidate for the MSSM Higgs-doublet pair.

The last term of eq. 4.14 is in the form of eq. 2.12, with two pairs of $\chi^c + \overline{\chi}^c$ fields. With $\langle S_{26}\rangle \neq 0$ an F-flat solution is found with $\langle\chi_1^c\rangle = \langle\overline{\chi}_1^c\rangle = 0$, then this part of the superpotential reduces to eq. 2.12. Employing F and D flatness conditions, the potential has a flat direction along the right-handed-neutrino direction of $\chi_2^c + \overline{\chi}_2^c$, as in eq. 2.10. The non-vanishing vevs break the PS to the SM gauge group and subsequently gives mass to the remaining massless color triplets in C_4 and the $D^c + \overline{D}^c$ components in $\chi_2^c + \overline{\chi}_2^c$. The remaining charged states in $\chi_2^c + \overline{\chi}_2^c$ obtain mass via the super-Higgs mechanism. Lastly, the $\chi_1^c + \overline{\chi}_1^c$ also acquire mass via the $\langle S_{26}\rangle$ vev in eq. 4.14.

The remaining central problem is whether the required singlets, such as those in eq. 4.13, could develop the appropriate vevs along F- and D-flat directions. There are five abelian factors in the A1 model, one of them is anomalous (which is cancelled by the generalized Green-Schwarz mechanism [42, 43], as usual). In general, the anomalous U_1 factor destabilizes the original vacua and contributes vevs to some of the singlet fields. It requires further investigation to determine whether our assumptions in the previous paragraph are substantiated, and moreover whether all the abelian symmetries may be broken with the vacuum solution. (The $\mathcal{N} = 1$ supersymmetry, however, is generically preserved in the effective theory.) Refs. [11, 44] have already obtained some

necessary conditions for analyzing non-trivial singlet vevs. We leave a detailed investigation for the future.

2. D_4 *family symmetry*

Before discussing the Yukawa matrices for quarks and leptons, we consider the family symmetry of model A1. The third family is a bulk field, while the first two families are located on the two $\mathbb{Z}_2$ fixed points in the SO_4 torus with an SO_{10} gauge symmetry. One family sits at each fixed point (see fig. 4). Since the Wilson line in the SO_4 torus lies in the orthogonal direction to these two fixed points, the theory is invariant under the permutation of the first two families, labelled by an index $n_2' = 0, 1$ (or $A = 1, 2$). In addition, the string selection rule, eq. C8, requires that every effective fermion mass operator include an even number of fields with $n_2' = 1$. Hence these effective operators are invariant under a $\mathbb{Z}_2$ parity $n_2' \to -n_2'$. The two operations are generated by the two Pauli matrices $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ acting on a real two dimensional vector. The complete set of operations closes on the discrete non-abelian family symmetry group $\mathrm{D}_4 = \{\pm I,\ \pm\sigma_1,\ \pm\sigma_3,\ \mp \mathrm{i}\sigma_2\}$. Note that the eight-element finite (dihedral) group D_4 is the symmetry group of a square. It has five conjugacy classes and five faithful representations. The character table is

Classes	I	$-I$	$\pm\sigma_1$	$\pm\sigma_3$	$\mp\mathrm{i}\sigma_2$
Doublet $-$ D	2	-2	0	0	0
Singlet $-$ A_1	1	1	1	1	1
Singlet $-$ B_1	1	1	1	-1	-1
Singlet $-$ B_2	1	1	-1	1	-1
Singlet $-$ A_2	1	1	-1	-1	1

. (4.17)

In our models, the first two families transform as the doublet, while the third family transforms as the trivial singlet.

We have many SO_{10} singlets in our models, transforming as doublets under D_4. They appear in effective higher dimension fermion mass operators. Consider, for example, two doublets under D_4 given by $\{S_A, \tilde{S}_A\}$. Then in terms of these two doublets we can define bilinear combinations transforming as $\{A_1, A_2, B_1, B_2\}$. We have

$$S_1\tilde{S}_1 + S_2\tilde{S}_2 \sim A_1$$

$$S_1\tilde{S}_2 - S_2\tilde{S}_1 \sim A_2$$

$$S_1\tilde{S}_2 + S_2\tilde{S}_1 \sim B_1$$
$$S_1\tilde{S}_1 - S_2\tilde{S}_2 \sim B_2 \tag{4.18}$$

The effective Yukawa couplings are then constructed in terms of D_4 invariants. Define the D_4 doublet left-handed quarks and leptons $(\mathbf{4},\mathbf{2},\mathbf{1})$ $[= f_A]$ and left-handed anti-quarks and anti-leptons $(\bar{\mathbf{4}},\mathbf{1},\mathbf{2})$ $[= f_A^c]$ for the first two families and the Higgs multiplet $(\mathbf{1},\mathbf{2},\mathbf{2})$ $[= h]$. We then have the PS and D_4 invariants:

$$hA_1(f_1 f_1^c + f_2 f_2^c) \equiv hA_1(f_A f_A^c)$$
$$hA_2(f_1 f_2^c - f_2 f_1^c)$$
$$hB_1(f_1 f_2^c + f_2 f_1^c)$$
$$hB_2(f_1 f_1^c - f_2 f_2^c) \tag{4.19}$$

We can also have operators of the form

$$h(f_A S_A)(f_B^c S_B) = h[f_1 f_1^c S_1^2 + f_2 f_2^c S_2^2 + (f_1 f_2^c + f_2 f_1^c) S_1 S_2] \tag{4.20}$$

Unfortunately there are, in principle, several possible ways of constructing D_4 invariants. We are not able to determine, without further string calculations, how to contract the D_4 indices. In the following we assume, for illustrative purposes, that only the simplest invariants, A_1 and B_1, appear in the effective Yukawa couplings.

3. Fermion masses

The only Yukawa coupling in model A1 present at leading order is for the third family, given by the first term in eq. C11. From discussions in sect. II B, we conclude that this coupling unifies with the GUT gauge coupling at the 5d compactification scale, as in eq. 2.7. Yukawa couplings for the first two families come from higher-dimensional non-renormalizable operators. In addition they are constrained by the D_4 family symmetry. In principle, there are at least two possible types of operators. The first type involves operators of the form $h_1(f_3 f_A^c,\ f_A f_3^c,\ f_A f_B^c)$, multiplied by suitable singlets, and the second type also involves composite singlets $\overline{\chi}^c\chi^c$. The second type of operators is particularly important since it has the potential to discriminate up-type quarks and charged leptons from the down-type; this is necessary to obtain a realistic CKM matrix and also resolve the "bad" GUT relation $m_s/m_d = m_\mu/m_e$.

We define the following two composite operators

$$\mathcal{O}_1 = \overline{\chi}_1^c\chi_\alpha^c, \qquad \mathcal{O}_2 = \overline{\chi}_2^c\chi_\alpha^c, \tag{4.21}$$

where the group indices are arranged in all possible ways. From the string selection rules, it is straightforward to show that the Yukawa matrix is (we only keep representative terms, see eq. C22 for more complete expressions),

$$(f_1\, f_2\, f_3)h_1 \begin{pmatrix} \mathcal{O}_2 S_{\rm e}^{(3,9,12)} + S_{\rm e}^{(10,22,22,23)} & \mathcal{O}_2 S_{\rm o}^{(3,9,12)} + S_{\rm o}^{(10,22,22,23)} & \mathcal{O}_1\mathcal{O}_2 S_{\rm e}^{(3,12)} + S_{\rm e}^{(9,10,22,22,23)} \\ \mathcal{O}_2 S_{\rm o}^{(3,9,12)} + S_{\rm o}^{(10,22,22,23)} & \mathcal{O}_2 S_{\rm e}^{(3,9,12)} + S_{\rm e}^{(10,22,22,23)} & \mathcal{O}_1\mathcal{O}_2 S_{\rm o}^{(3,12)} + S_{\rm o}^{(9,10,22,22,23)} \\ S_{\rm e}^{(10,26)} & S_{\rm o}^{(10,26)} & 1 \end{pmatrix} \begin{pmatrix} f_1^c \\ f_2^c \\ f_3^c \end{pmatrix}, \tag{4.22}$$

where

$$S_{\rm e}^{(a,b,\cdots)} = \sum_{\sum A = \rm even} S_A^a S_B^b \cdots, \qquad S_{\rm o}^{(a,b,\cdots)} = \sum_{\sum A = \rm odd} S_A^a S_B^b \cdots, \tag{4.23}$$

with A's are the family indices of the corresponding singlets (of the $T_{1,3}$ sectors). Several comments on eq. 4.22 are now in order.

- The structure of the Yukawa matrix is determined by a D_4 family symmetry. (See the caveat at the end of Section IV B 2.)

- Given the superpotential for color triplets, eq. 4.14, an F-flat direction requires $\langle \overline{\chi}_1^c \rangle = 0$ which gives $\mathcal{O}_1 = 0$. If however there is a higher-dimensional operator of the form $S_{26} S \chi_\alpha^c \overline{\chi}_2^c$, then the combined terms $S_{26}(\chi_\alpha^c \overline{\chi}_1^c + S \chi_\alpha^c \overline{\chi}_2^c)$ has an F-flat solution with $\langle \overline{\chi}_1^c \rangle, \langle \overline{\chi}_2^c \rangle \neq 0$ and thus $\mathcal{O}_1, \mathcal{O}_2 \neq 0$. We will analyze the more general case.

- Given the superpotential for color triplet masses, our previous solution eq. 4.16 requires $\langle S_9 \rangle = 0$. Hence the composite operators $S^{(3,9,12)}$, $S^{(9,10,22,22,23)}$ vanish. However, it is again possible that these terms may still be present when higher order products of operators are considered.

- It is crucial to understand how the PS group indices are contracted and what the corresponding Clebsch-Gordon (CG) coefficients are. In orbifold models, the massless matter fields correspond to the (integral) highest weight representations of the level-one Kač-Moody algebra. In principle one may extract the desired information from the conformal blocks. We have not attempted to perform such a string theoretical analysis. Instead we shall adopt a simpler field theoretical approach, following ref. [45].

Our aim is to examine phenomenological implications of eq. 4.22 in a simple setting. We shall consider two simple cases in the following.

Case A –

First neglect the (13) and (23) entries of eq. 4.22 (it may be reasonable to do so, because they are higher order terms), and consider the 2×2 sub-matrix corresponding to the second and third families, i.e.

$$\begin{pmatrix} \mathcal{O}_2 S_{\rm e}^{(3,9,12)} + S_{\rm e}^{(10,22,22,23)} & 0 \\ S_{\rm o}^{(10,26)} & 1 \end{pmatrix}. \tag{4.24}$$

We may take $\mathcal{O}_2$ to be in the form of the $\mathcal{O}^W$ operator of ref. [45]. This operator has a vanishing (non-vanishing) CG coefficient for the up (down) type fields. One may require

$$\langle \mathcal{O}_2 S_{\rm e}^{(3,9,12)} \rangle \sim \frac{m_s}{m_b} \sim \lambda^2, \qquad \langle S_{\rm e}^{(10,22,22,23)} \rangle \sim \frac{m_c}{m_t} \sim \lambda^3, \tag{4.25}$$

where $\lambda \simeq 0.22$ is the Cabibbo angle. The mixing angle V_{cb} is then approximately $\lambda^2 \langle S_{\rm o}^{(10,26)} \rangle$, implying $S_{\rm o}^{(10,26)} \sim \mathcal{O}(1)$.

Next consider the (2×2) sub-matrix corresponding to the first and second families. Note that this part always has the following form,

$$\begin{pmatrix} a_{u,d} & b_{u,d} \\ b_{u,d} & a_{u,d} \end{pmatrix}. \tag{4.26}$$

The democratic form may lead to realistic values for quark masses and mixings. Taking $a_{u,d} = b_{u,d}(1 + \varepsilon_{u,d})$ with $\varepsilon_d \approx \lambda$ (which implies an approximate $\mathbb{Z}_2$ symmetry between vevs, $\mathcal{O}_2 S_{\rm e}^{(3,9,12)} = \mathcal{O}_2 S_{\rm o}^{(3,9,12)}(1 + \varepsilon_d)$, $S_{\rm e}^{(10,22,22,23)} = S_{\rm o}^{(10,22,22,23)}(1 + \varepsilon_u)$ and $S_{\rm e}^{(10,26)} \simeq S_{\rm o}^{(10,26)}$), we obtain the mass ratio m_d/m_s and CKM angle V_{us} at correct orders. More suppressed value for ε_u, e.g. $\varepsilon_u \sim 10^{-3}$, is required for m_u/m_c. Finally, it is also possible to obtain correct mass relations for the charged leptons, $m_\mu/m_\tau \sim m_s/m_b$ and $m_e/m_\mu \sim m_d/m_s$.

Case B –

Consider now the case that the (13) and (23) entries are not negligible. Let us parameterize the (23) entry in the down sector by y_{23}^d and assume the corresponding entry in the up sector is smaller (or comparable). For simplicity, we also assume $S^{(10,26)} = 0$. The (2×2) sub-matrix of the second and third families is

$$\begin{pmatrix} \mathcal{O}_2 S_{\rm e}^{(3,9,12)} + S_{\rm e}^{(10,22,22,23)} & y_{23}^d \\ 0 & 1 \end{pmatrix}. \tag{4.27}$$

We can obtain correct mass ratios m_c/m_t and m_s/m_b and mixing angle V_{cb} if $y_{23}^d \sim \mathcal{O}_2 S_{\rm e}^{(3,9,12)} \sim \lambda^2$ and $S_{\rm e}^{(10,22,22,23)} \sim \lambda^3$. The discussion on the first and second families follows essentially in the

same way as in case A. However, we now need to tune $\mathcal{O}_2 S^{(3,9,12)} + S^{(10,22,22,23)}$ appropriately.

Finally consider neutrino masses. The Dirac neutrino mass matrix has the same form as that of eq. C22. Effective Majorana neutrino masses are obtained in eq. C24, where the non-trivial effective operators have the form $f^c_a f^c_b \overline{\chi}^c_i \overline{\chi}^c_j$ $(a, b = 1, 2, 3,\ i, j = 1, 2)$ with suitable powers of singlets. The non-vanishing vevs of $\overline{\chi}^c_i$ project out the right-handed neutrinos in $f^c_3,\ f^c_A$. One then obtains a Majorana mass of order $< M^2_{\rm PS}/M_{\rm string} \simeq M_{\rm GUT}/(2{\rm e}^3) \simeq 7 \times 10^{14}$ GeV, which is just right for generating acceptable light neutrino masses via the see-saw mechanism. Although at the present operator order the Majorana mass terms vanish with the non-vanishing vevs discussed earlier, non-trivial operators may exist at higher order.

V. CONCLUSION

In this paper we construct three-family PS models in the $\mathbb{Z}_6$ abelian symmetric orbifold. Our models are mainly motivated by recent discussions on orbifold GUTs. We are able to realize some features of the orbifold GUTs in the string compactification limit where the compactified space is effectively 5d. The breaking of the $\mathcal{N} = 2$ to $\mathcal{N} = 1$ supersymmetry in 4d and the $\rm E_6$ (or $\rm SO_{10}$) gauge symmetry to that of PS are realized similarly as in orbifold GUTs. We find three family chiral matter fields, two of them can be regarded as “brane” states and one as a “bulk” state, in the terminology of orbifold GUTs. These models extend three-family orbifold string model building to non-prime-order orbifolds, and we find some new features in the matter spectra when compared to that of the prime-order orbifolds. Matter fields arise not only from the untwisted but also from twisted sectors, and typically there is a horizontal $2 + 1$ splitting in the family space. This splitting may have the potential to better facilitate the description of fermion masses and mixings.

We find one of our string models, with an $\rm E_6$ gauge group in 5d, is particularly interesting. It has the following properties:

1. Renormalizable Yukawa couplings exist only for the third family. Moreover the model predicts a unification relation among the third family Yukawa couplings and GUT gauge coupling at the 5d compactification scale.

2. The renormalizable and non-renormalizable couplings can affect a spontaneous breaking of the PS symmetry to that of the SM with fields in the $(\mathbf{6}, \mathbf{1}, \mathbf{1})$ and $(\mathbf{4}, \mathbf{1}, \mathbf{2}) + (\overline{\mathbf{4}}, \mathbf{1}, \mathbf{2})$ representations. Moreover, after the symmetry breaking, many unwanted states can develop large mass.

3. There is a non-abelian D_4 family symmetry for the first two families, which makes the model a good playground for studying fermion mass hierarchy and the flavor problem in supersymmetry.

We regard these phenomenological features as merits of our models.

On the other hand, there are two main problems for the models (of course, they are not unique to our models).

- Exotics: There is a pair of SM vector-like exotic particles at the low energy scale. To the order of our analysis, we have not found any Yukawa coupling that can give them mass.
- Proton stability: We have Higgs fields transforming in the same (or conjugate) PS representation as the SM right-handed matter. In general, baryon/lepton-number violating effective operators are induced after PS symmetry breaking. We have not yet found a symmetry that can distinguish the PS breaking Higgses from matter and effectively eliminate the dangerous operators to all orders.

These problems have only been examined at a rather primitive and qualitative level. Specifically we have looked for non-trivial Yukawa couplings allowed by string selection rules to certain orders. We have shown that the baryon/lepton-number violating effective operators can be avoided if one prudently chooses appropriate values for the singlet vevs. Unfortunately, we are not able to extend this argument to all orders in string perturbation. The problem apparently results from the presence of *only the trivial fixed point* in the T_1 twisted sector on the G_2 torus. This has the consequence of nullifying any space group selection rules for the G_2 torus whenever any T_1 twisted sector operator is present.

In any case, these problems may point to the direction for refining our models. For example, it would be desirable to search for models with fewer color triplets and, if possible, no exotics. The analysis presented here is clearly just the beginning. It would certainly be useful to expand the search to other $\mathbb{Z}_2 \times \mathbb{Z}_N$ orbifolds with one (or two) Wilson lines in the SO_4 direction in order to find more effective 5 or 6d orbifold GUTs. (The $\mathbb{Z}_2 \times \mathbb{Z}_2$ model might be particularly interesting and it has been studied recently in ref. [22]. Presumably the model is simpler than ours because it has only three twisted sectors and the number of modular invariant gauge embeddings is more limited. However, realistic three-family model have yet to be constructed.) In brief, we believe our analysis has opened up a promising new direction for string model building.

Acknowledgments

This work has been presented at the Planck '04 and String Phenomenology '04 conferences and the Aspen "String and Real World" workshop. We would like to thank the participants of these meetings, especially A. Faraggi, H. P. Nilles and S. Stieberger, for stimulating conversations. T. K. was supported in part by the Grant-in-Aid for Scientific Research (#16028211) and the Grant-in-Aid for the 21st Century COE "The Center for Diversity and Universality in Physics" from Ministry of Education, Science, Sports and Culture of Japan. S. R. and R.-J. Z. were supported in part by DOE grants DOE/ER/01545-858 and DE-FG02-95ER40893 respectively. They also thank the Aspen Center for Physics for hospitality during the final stage of this work.

APPENDIX A: RECIPES FOR CONSTRUCTING NON-PRIME-ORDER ORBIFOLD MODELS IN THE PRESENCE OF DISCRETE WILSON LINES

In this appendix we review the construction of non-prime-order orbifold models with discrete Wilson lines.

Our starting point is the 10d heterotic string theory, which consists of a 26d left-moving bosonic string and a 10d right-moving superstring. Modular invariance requires the momenta of the internal left-moving bosonic degrees of freedom (16 of them) lie in a 16d Euclidean even self-dual lattice, we choose to be the $\mathrm{E}_8 \times \mathrm{E}_8$ root lattice.[12]

To make a connection to the 4d world, we must compactify 6 spatial dimensions of the 10 remaining space-time dimensions. There are many ways (see, e.g., ref. [5] for a collection of early works) to achieve a 4d $\mathcal{N} = 1$ supersymmetric spectrum, among them the most studied in the literature is the orbifold construction [4, 6, 9, 10, 11, 24]. We will use the simplest abelian symmetric orbifold construction, where we differentiate the space-time and internal degrees of freedom, and realize the orbifold twists and Wilson lines by shifts in the $\mathrm{E}_8 \times \mathrm{E}_8$ lattice. This type of construction admits a clear space-time interpretation.

The full definition of an orbifold model requires the specification of a six-torus T^6, corresponding to the compactified spatial dimensions, a point group, (such as the cyclic groups $\mathbb{Z}_N$ or $\mathbb{Z}_N \times \mathbb{Z}_M$ with $N,\ M = 3, 4, 6, 7, 8, 12$ [4]), corresponding to the automorphism of the T^6-lattice, and an embedding of the space group, (which consists of the point group and the lattice translations),

[12] For an orthonormal basis, the E_8 root lattice consists of following vectors, $(n_1, n_2, \cdots, n_8)$ and $(n_1 + \frac{1}{2}, n_2 + \frac{1}{2}, \cdots, n_8 + \frac{1}{2})$, where $n_1, n_2, \cdots n_8$ are integers and $\sum_{i=1}^{8} n_i = 0 \bmod 2$.

in the $\mathrm{E}_8 \times \mathrm{E}_8$ lattice. Normally one denotes the generator of the discrete group $\mathbb{Z}_N$ by a *twist vector* $\mathbf{v} = (v_1, v_2, v_3)$, acting on the three complex planes by $\theta : Z_i \to \mathrm{e}^{2\pi \mathrm{i} v_i} Z_i$ $(i = 1, 2, 3)$. To ensure that one space-time supersymmetry survives in 4d, the twist vector needs to satisfy $\pm v_1 \pm v_2 \pm v_3 = 0 \bmod 2$,[13] and none of the v_i's vanishes. In the abelian orbifold construction, embeddings of the space group in the $\mathrm{E}_8 \times \mathrm{E}_8$ are realized by shifts of the corresponding lattice, $\mathbf{P} \to \mathbf{P} + k\mathbf{V} + l\mathbf{W}$, where k, l are integers, $\mathbf{P}$ are vectors of the $\mathrm{E}_8 \times \mathrm{E}_8$ root lattice, $\mathbf{V}$ the *gauge twists*, realizing the point group, and $\mathbf{W}$ the *discrete Wilson lines*, realizing the lattice translations. The cyclic group multiplication rules require $N\mathbf{V}$ and $N_W\mathbf{W}$ to be in the $\mathrm{E}_8 \times \mathrm{E}_8$ lattice. (In general the degrees of the Wilson lines, N_W, divide the degree of the orbifold twist, N.)

String states closed on T^6, i.e. those satisfying the condition $Z_i(\tau, \sigma + \pi) = Z_i(\tau, \sigma)$ modulo lattice translations, give rise to the untwisted-sector states. Besides the $\mathcal{N} = 1$ supergravity multiplet and modulus fields that parameterize deformations of the background fields, the untwisted-sector states also give rise to gauge and matter fields. Embeddings of the point group in $\mathrm{E}_8 \times \mathrm{E}_8$ break the gauge symmetry down to its commutator subgroups, i.e., the surviving non-zero roots satisfy

$$\mathbf{P}^2 = 2\,, \qquad \mathbf{P} \cdot \mathbf{V} \in \mathbb{Z}\,. \tag{A1}$$

Note that we cannot lower the rank of the surviving groups in abelian orbifold models, since by construction the gauge twists and Wilson lines commute with the $\mathrm{E}_8 \times \mathrm{E}_8$ Cartan subalgebra.

The conditions for the untwisted-sector matter states are similar. It is convenient to bosonize the right-moving fermionic degrees of freedom and denote their SO_8 momenta in the light-cone gauge by $\mathbf{r}$.[14] The right-moving NS sectors ($b^i_{-1/2}$ for $i = 1, 2, 3$) of the untwisted matter states have SO_8 weights $\mathbf{r}_1 = (0, 1, 0, 0)$, $\mathbf{r}_2 = (0, 0, 1, 0)$ and $\mathbf{r}_3 = (0, 0, 0, 1)$ and they pick up phases $\mathrm{e}^{-2\pi \mathrm{i} \mathbf{r}_i \cdot \mathbf{v}}$ under the orbifold twists, so the corresponding roots must satisfy

$$\mathbf{P}^2 = 2\,, \qquad \mathbf{P} \cdot \mathbf{V} - \mathbf{r}_i \cdot \mathbf{v} \in \mathbb{Z} \qquad (i = 1, 2, 3)\,. \tag{A2}$$

This gives three untwisted-matter sectors U_i, one for each complex plane. (Eq. A1 can also be written in the same form with $\mathbf{r}_0 = (1, 0, 0, 0)$, where the non-zero entry lies in the uncompactified direction).

The gauge groups and untwisted-sector matter spectra are modified when discrete Wilson lines are turned on. In the presence of general background fields G_{ij}, B_{ij} and $\mathbf{W}_i$, the canonical momenta

[13] The signs are arbitrary. We will use the convention that all signs are positive.

[14] That is, $\mathbf{r}$'s are in the SO_8 weight lattice, the integral and half-integral weights correspond to the Neveu-Schwarz (NS) and Ramond (R) sector fermions. For R-sector weights, the first components, r_0, indicate the helicities of space-time fermions. In this notation, the first component of the twist vector, v_0, is zero. $\mathbf{r}$ are commonly referred to as the H momenta.

conjugate to the compactified coordinates and the gauge coordinates are $\Pi^i = p_L^i + p_R^i + \frac{1}{2}\mathbf{W}^i \cdot (\mathbf{\Pi}+\mathbf{p}_L) + B^{ij}(p_{Lj} - p_{Rj})$ and $\mathbf{\Pi} = \mathbf{p}_L - \frac{1}{2}\mathbf{W}_i(p_L^i - p_R^i)$. Since $p_L^i - p_R^i = 2n^i$, $\Pi^i = m^i$ and $\mathbf{\Pi} = \mathbf{P}$, where the integers m^i, n^i are the momentum (or KK modes) and winding quantum numbers, we have [25, 46] (the string unit is $\alpha' = 1/2$)

$$\begin{aligned} p_L^i &= \frac{m^i}{2} + \left(G^{ij} - B^{ij} - \frac{1}{4}\mathbf{W}^i \cdot \mathbf{W}^j\right) n_j - \frac{1}{2}\mathbf{P}\cdot\mathbf{W}^i \,, \\ p_R^i &= p_L^i - 2n^i \,, \qquad \mathbf{p}_L \;=\; \mathbf{P} + n^i \mathbf{W}_i \,. \end{aligned} \tag{A3}$$

In the models studied in this paper, we take $B_{ij} = 0$ and $G_{ij} = \frac{1}{2}R_i R_j \mathbf{e}_i \cdot \mathbf{e}_j$ is defined by the geometry of the internal space T^6 with basis vectors given by $\mathbf{e}_i$ and dimensions by R_i. In this case, these equations, along with the masslessness conditions

$$\frac{1}{4}m_R^2 = N_R + \frac{1}{2}G_{ij}p_R^i p_R^j + \frac{1}{2}\mathbf{r}^2 - \frac{1}{2} = 0 \tag{A4}$$

$$\frac{1}{4}m_L^2 = N_L + \frac{1}{2}G_{ij}p_L^i p_L^j + \frac{1}{2}\mathbf{p}_L^2 - 1 = 0 \tag{A5}$$

where $N_L,\ N_R$ are integral oscillator mode numbers and the last two terms in eq. A4 are the contribution of the bosonized NSR fermions, require the winding number for massless states to be zero, and the last term of p_L^i an integer (for both gauge and matter fields), i.e.,

$$\mathbf{P}\cdot\mathbf{W} \in \mathbb{Z} \,. \tag{A6}$$

String states closed on themselves under the identification of a non-trivial element of the point group (modulo translations by lattice vectors) give rise to the twisted-sector states. For the k^{th} twisted-sector T_k (for which the complex compactified coordinates satisfy $Z_i(\tau, \sigma+\pi) = \mathrm{e}^{2\pi \mathrm{i} k v_i} Z_i(\tau,\sigma)$), the $\mathrm{E}_8 \times \mathrm{E}_8$ and SO_8 momenta are shifted according to $\mathbf{P} \to \mathbf{P} + k\mathbf{X}_{n_\mathbf{f}}$ and $\mathbf{r} \to \mathbf{r} + k\mathbf{v}$, where $\mathbf{X}_{n_\mathbf{f}} = \mathbf{V} + n_\mathbf{f}^i \mathbf{W}_i$ with $n_\mathbf{f}^i$ being fixed-point dependent winding numbers.[15] The massless states satisfy the following equations [4, 9, 10, 11, 24],

$$\frac{1}{4}m_R^2 = \sum_{i=1}^{6} N_i^R \omega_i^{(k)} + \frac{1}{2}\left(\mathbf{r} + k\mathbf{v}\right)^2 + a_R^{(k)} = 0 \,, \tag{A7}$$

$$\frac{1}{4}m_L^2 = \sum_{i=1}^{6} N_i^L \omega_i^{(k)} + \frac{1}{2}\left(\mathbf{P} + k\mathbf{X}_{n_\mathbf{f}}\right)^2 + a_L^{(k)} = 0 \,, \tag{A8}$$

where N_i^R and N_i^L are intergral numbers of the right- and left-moving (bosonic) oscillators, $a_R^{(k)}$, $a_L^{(k)}$ are the normal ordering constants,

$$a_R^{(k)} \;=\; -\frac{1}{2} + \frac{1}{2}\sum_{i=1}^{3} |\widehat{kv_i}| \left(1 - |\widehat{kv_i}|\right) \,,$$

[15] Discrete Wilson lines break the degeneracies of the fixed points. The integers $n_\mathbf{f}^i$ should be chosen appropriately, depending on the direction and degree of the Wilson line [25, 47, 48].

$$a_L^{(k)} \;=\; -1+\frac{1}{2}\sum_{i=1}^{3}|\widehat{kv_i}|\left(1-|\widehat{kv_i}|\right)\,, \tag{A9}$$

with $\widehat{kv_i}=kv_i \bmod 1$, and $\omega_i^{(k)}=\widehat{kv_i}$ if $\widehat{kv_i}>0$ and $1-\widehat{kv_i}$ if $\widehat{kv_i}\le 0$ are oscillator energies. Thus, a twisted-sector matter state can be labelled by the twisted sector, k, the fixed point it is localized on, $\mathbf{f}$, the number of windings associated with the fixed point, $n^i_{\mathbf{f}}$, and the number of right- and left-moving oscillators, N_i^R and N_i^L.

Note, however, that not all gauge twists and discrete Wilson lines are physically allowed. To ensure modular invariance of the one-loop partition function (or the level-matching condition) for the right- and left-movers, one needs to require [4, 9, 10, 11, 24, 26]

$$N(\mathbf{X}^2_{n_{\mathbf{f}}}-\mathbf{v}^2)=0 \bmod 2\,. \tag{A10}$$

In addition, in non-prime-order orbifold models, the degrees of Wilson lines are in general the divisors of that of the orbifold twist; they need to satisfy more stringent modular-invariance requirements. For example, in the $\mathbb{Z}_6$ model where $\mathbf{v}_6=\frac{1}{6}(1,2,-3)$ there are at most three admissible Wilson lines, two of degree-2 ($\mathbf{W}_2^{(i)}$, $i=1,2$) and one of degree-3 ($\mathbf{W}_3$) [47, 48].[16] Modular-invariance conditions involving these Wilson lines are the following more restrictive set (where $i=1,2$),

$$\{2(\mathbf{W}_2^{(i)})^2\,,\ 3(\mathbf{W}_3)^2\,,\ 4\mathbf{W}_2^{(1)}\cdot\mathbf{W}_2^{(2)}\,,\ 12\mathbf{W}_2^{(i)}\cdot\mathbf{W}_3\}=0 \bmod 2. \tag{A11}$$

Further complications arise for non-prime-order orbifold models. Unlike the prime-order orbifolds such as the $\mathbb{Z}_3$ and $\mathbb{Z}_7$ models, fixed points of the higher-twisted sectors are not always invariant under the defining orbifold twist, θ. However, one can find linear combinations of the states corresponding to these fixed points such that they have definite eigenvalues under the θ rotation. As it has been shown in ref. [48, 49], for any fixed point represented by a space-group element $(\theta^k,\mathbf{l})$ (i.e., a fixed point $\mathbf{f}$ satisfying $\mathbf{f}=\theta^k\mathbf{f}+\mathbf{l}$, where $\mathbf{l}$ is a vector of the T^6 lattice and called the *equivalent shift vector*), if it can be written as a power of a prime element $(\theta^m,\mathbf{l}')$ (a

[16] Actually, the number of admissible Wilson lines also depends on the compactified lattice. If we restrict it to a Lie algebra root lattice, then there are four possibilities, (A) $\mathrm{SU}_6\oplus\mathrm{SU}_2$, (B) $\mathrm{SU}_3\oplus\mathrm{SO}_8$, (C) $\mathrm{SU}_3\oplus\mathrm{SO}_7\oplus\mathrm{SU}_2$, and (D) $\mathrm{G}_2\oplus\mathrm{SU}_3\oplus\mathrm{SO}_4$, whose Coxeter elements realize the $\mathbb{Z}_6$ orbifolding. (Note that the lattices C and D are the sublattices of B, where the quotients are isomorphic to $\mathbb{Z}_2$ and $\mathbb{Z}_2\times\mathbb{Z}_2$. The $\mathbb{Z}_2\times\mathbb{Z}_2$ is the center of SO_8. There are two additional lattices, $\mathrm{SU}_4{}^{[2]}\oplus\mathrm{SU}_3\oplus\mathrm{SU}_2$ and $\mathrm{SU}_3{}^{[2]}\oplus\mathrm{SU}_3\oplus\mathrm{SO}_4$, which also involve outer automorphism of the root lattice and give identical results to C and D.) The A (B/C) lattice has at most one degree-2 (one degree-2 and one degree-3) Wilson line(s). The D lattice corresponds to the case in the main text and is the most intuitive one. For our three-family models, we choose to turn on two Wilson lines, one degree 2 and one degree 3. The low energy phenomenology for the last three lattices are different, since both the quantum numbers (which label the states) and the selection rules (which determine allowed couplings) depend on the lattice choice. In particular, the D_4 family symmetry does not subsist for the B and C lattices.

prime element is the one that cannot be written as a power of any other element), then the linear combinations,

$$|k,\gamma\rangle = \sum_{\ell=0}^{m-1} \gamma^{-\ell} |\theta^k, \theta^\ell \mathbf{l}\rangle \,, \tag{A12}$$

have eigenvalues $\gamma = \mathrm{e}^{2\pi i n/m}$ and $n = 0, 1, \cdots, m-1$ under the θ rotation. We can therefore use this notation to appropriately label twisted-sector states arising from different fixed points according to their θ-eigenvalues.

Finally, in the $\mathbb{Z}_6$ orbifold that we are most interested in (and other orbifolds containing sub-orbifolds with fixed tori), special care must also be taken in the presence of discrete Wilson lines. This orbifold contains $\mathbb{Z}_2$ and $\mathbb{Z}_3$ sub-orbifolds with twists $\{1, \theta^3\}$ and $\{1, \theta^2, \theta^4\}$ respectively. The second and third complex planes are invariant tori of the θ^3 and θ^2/θ^4 twists, i.e., they are unrotated under the respective orbifold twists. Consequently these directions have mode expansions of a toroidal coordinate in the T_3 and $T_{2,4}$ sectors, as in eq. A3, with gauge momenta replaced by the shifted momenta, $\mathbf{P}+3\mathbf{X}_{n_\mathbf{f}}$ in the T_3 and $\mathbf{P}+2\mathbf{X}_{n_\mathbf{f}}$, $\mathbf{P}+4\mathbf{X}_{n_\mathbf{f}}$ in the T_2, T_4 sectors. There are also Wilson-line dependent terms. The mode expansions of the second direction in the T_3 sector and the third direction in the $T_{2,4}$ sectors contain the $2\mathbf{W}_3$ and $3\mathbf{W}_2^{(i)}$ $(i = 1, 2)$ Wilson lines respectively. (The numeric factors are due to the fact that the only non-trivial action for the $\mathbf{W}_3$ and $\mathbf{W}_2^{(i)}$ Wilson lines are the $\mathbb{Z}_2$ and $\mathbb{Z}_3$ sub-orbifold twists.) Just like the untwisted-sector states, the winding numbers for massless states along these unrotated directions in respective twisted sectors must be set to zero, and the shifted momenta must satisfy the (necessary) projection conditions $2(\mathbf{P}+3\mathbf{X}_{n_\mathbf{f}})\cdot\mathbf{W}_3 \in \mathbb{Z}$ for the T_3 and $3(\mathbf{P}+2\mathbf{X}_{n_\mathbf{f}})\cdot\mathbf{W}_2^{(i)} \in \mathbb{Z}$, $3(\mathbf{P}+4\mathbf{X}_{n_\mathbf{f}})\cdot\mathbf{W}_2^{(i)} \in \mathbb{Z}$ for the T_2, T_4 sector states. Using the modular invariance conditions, eqs. A10 and A11, these conditions reduce to

$$\mathbf{P}\cdot\mathbf{W}_3 \in \mathbb{Z}\,, \quad \text{for } T_3 \text{ sector}\,; \qquad \mathbf{P}\cdot\mathbf{W}_2^{(i)} \in \mathbb{Z}\,, \quad \text{for } T_{2,4} \text{ sector}\,. \tag{A13}$$

These projections are crucial for rendering an anomaly-free mass spectrum and have not been properly addressed in the literature.

Multiplicities of the twisted-sector states are computed from the generalized Gliozzi-Scherk-Olive (GSO) projector [10, 28], which can be derived from the one-loop partition function by power-series expansions. In our notation, for a twisted-sector state labelled by $\{k, \gamma, n_\mathbf{f}\}$ and appropriate oscillator numbers, the projector is [47]

$$P(k,\gamma,n_\mathbf{f}) = \frac{1}{N}\sum_{\ell=0}^{N-1} \left[\Delta(k,\gamma,n_\mathbf{f})\right]^\ell \,, \tag{A14}$$

where

$$\Delta(k,\gamma,n_{\mathbf{f}}) = \phi\gamma \exp\left\{ \mathrm{i}\pi \left[(2\mathbf{P} + k\mathbf{X}_{n_{\mathbf{f}}}) \cdot \mathbf{X}_{n_{\mathbf{f}}} - (2\mathbf{r} + k\mathbf{v}) \cdot \mathbf{v} \right] \right\} . \tag{A15}$$

In this expression, $\phi = \exp[2\pi\mathrm{i}\sum_{i=1}^{6}(N_i^L - N_i^R)\hat{v}_i]$ is the oscillator phase, with $\hat{v}_i = \mathrm{sgn}(\widehat{kv_i})v_i$, $\hat{v}_{i+3} = -\mathrm{sgn}(\widehat{kv_i})v_i$ if $\widehat{kv_i} \neq 0$, and $\hat{v}_i, \hat{v}_{i+3} = 0$ if $\widehat{kv_i} = 0$, for $i = 1,2,3$. Only states with $\Delta(k,\gamma,n_{\mathbf{f}}) = 1$ will survive the projection. (Equivalently, the generalized GSO projector can be defined as in ref. [10], in a slightly different fashion. We find it is more convenient to implement eq. A14 in the non-prime-order orbifolds.)

The GSO projector in eqs. A14 and A15 needs to be modified slightly for states in the untwisted sectors and those twisted sectors with fixed tori. In the untwisted case where $k = 0$, $\gamma = \phi = 1$, $n_{\mathbf{f}} = 0$, the required modification is simply an additional projection with respect to the Wilson line,

$$\Delta^{\ell} \to \frac{1}{N_W} \sum_{m_{\mathbf{f}}=0}^{N_W-1} \exp\left[\mathrm{i}\pi\ell(2\mathbf{P} \cdot \mathbf{X}_{m_{\mathbf{f}}} - 2\mathbf{r} \cdot \mathbf{v}) \right] . \tag{A16}$$

Then $\Delta = 1$ gives $\mathbf{P} \cdot \mathbf{V} - \mathbf{r} \cdot \mathbf{v} \in \mathbb{Z}$, $\mathbf{P} \cdot \mathbf{W} \in \mathbb{Z}$, the same projection conditions as in eqs. A1, A2 and A6.

The second case is more complicated. Assume that the T_k sector has a fixed torus, and the Wilson line of degree N_W also lies in this torus. Then we must have N_W divides k, and the required modification to the GSO projector is

$$\Delta^{\ell} \to \frac{(\phi\gamma)^{\ell}}{N_W} \sum_{m_{\mathbf{f}}=0}^{N_W-1} \exp\left\{ \mathrm{i}\pi\ell \left[(2\mathbf{P} + k\mathbf{X}_{m_{\mathbf{f}}}) \cdot \mathbf{X}_{m_{\mathbf{f}}} - (2\mathbf{r} + k\mathbf{v}) \cdot \mathbf{v} \right] \right\} . \tag{A17}$$

In this expression, it must be understood that each shifted momentum, $\mathbf{P} + k\mathbf{X}_{m_{\mathbf{f}}}$, satisfies its own masslessness condition, eq. A8. Therefore the momenta $\mathbf{P}$ in different terms differ by a multiple of $k\mathbf{W}$, which is a vector in the $\mathrm{E}_8 \times \mathrm{E}_8$ root lattice; these states are isomorphic to each other. Requiring that $\Delta = 1$ in eq. A17 then implies an additional projection $\mathbf{P} \cdot \mathbf{W} \in \mathbb{Z}$ for the T_k twisted sector states. It can also be seen that with this condition eq. A17 reduces to eqs. A14 and A15 with $\mathbf{X}_{n_{\mathbf{f}}} = \mathbf{V}$. These discussions can be easily generalized to the cases with several Wilson lines. In the $\mathbb{Z}_6$ model, they give the same projection conditions as in eq. A13.

While it is quite tedious to carry out all these procedures in practice, they can be easily computerized, as follows. First we check the modular invariance of the gauge twists and Wilson lines, and find all the $\mathrm{E}_8 \times \mathrm{E}_8$ roots that satisfy eqs. A1 and A6. From these we find the set of linearly independent positive roots (i.e., the simple roots) and determine the unbroken gauge groups. Next

we find the roots corresponding to the untwisted- and twisted-sector matter states, from eqs. A2, A6, and eqs. A7, A8 respectively. For the twisted-sector matter, we only keep states of negative helicity (i.e. left-handed), and compute their multiplicities using the generalized GSO projector in eq. A14. We also need to take special care of the twisted-sector states in the presence of Wilson lines, as noted in the paragraph after eq. A12. We then decompose the roots into highest weight representations, and determine their Dynkin indices and dimensions using the standard group theoretical method. Finally, the remaining abelian factors are found by searching for orthogonal roots, and the charges for matter states are determined accordingly by projecting the matter states onto these abelian roots.

APPENDIX B: THREE-FAMILY PATI-SALAM MODELS

1. The models

In this appendix, we apply the procedure outlined in appendix A to construct three-family PS models in the $\mathbb{Z}_6$ orbifold with $\mathbf{v}_6 = \frac{1}{6}(1, 2, -3)$. The $\mathbb{Z}_6$ is equivalent to the $\mathbb{Z}_2 \times \mathbb{Z}_3$ orbifold, where the two twist vectors are $\mathbf{v}_2 = 3\mathbf{v}_6 = \frac{1}{2}(1, 0, -1)$ and $\mathbf{v}_3 = 2\mathbf{v}_6 = \frac{1}{3}(1, -1, 0)$.

There are in total 61 inequivalent modular invariant choices for the gauge twists in the $\mathbb{Z}_6$ orbifold model [24]. To narrow down the possibilities, we demand the models we start with (before imposing any Wilson line) contain an SO_{10} gauge group and some matter fields in $\mathbf{16}/\overline{\mathbf{16}}$ representations in the first or third twisted sectors. Although this step makes our results less generic, it greatly reduces the large number of possible models to a manageable subset. We choose the following two gauge twists,

• model A:

$$\mathbf{V}_6 = \frac{1}{6}\,(22200000)\,(11000000)\ , \tag{B1}$$

• model B:

$$\mathbf{V}_6 = \frac{1}{6}\,(41100000)\,(22000000)\ , \tag{B2}$$

which break the $\mathrm{E}_8 \times \mathrm{E}_8$ gauge symmetry down to $\mathrm{SO}_{10} \times \mathrm{SU}_3 \times \mathrm{E}_7{}'$ and $\mathrm{SO}_{10} \times \mathrm{SU}_2 \times \mathrm{E}_7{}'$ respectively. The model A (B) contains four (two) $\mathbf{16}$ and one (three) $\overline{\mathbf{16}}$ in the untwisted sectors, and eighteen (fourteen) $\mathbf{16}$ and three (seven) $\overline{\mathbf{16}}$ in the twisted sectors; in total there are eighteen (six) SO_{10} families [24].

To further break the gauge symmetries and reduce the number of families, we impose discrete Wilson lines. As previously mentioned, there are at most one degree-3 Wilson line in the second complex plane, and two degree-2 Wilson lines in the third. We choose to add two of them, one of degree-2 and one of degree-3, as follows,

• model A1:

$$\mathbf{W}_2 = \frac{1}{2}\,(10000111)\,(00000000)\ , \qquad \mathbf{W}_3 = \frac{1}{3}\,(1-1000000)\,(00200000)\ , \tag{B3}$$

• model A2:

$$\mathbf{W}_2 = \frac{1}{2}\,(10000111)\,(00000000)\ , \qquad \mathbf{W}_3 = \frac{1}{3}\,(21-100000)\,(02110000)\ , \tag{B4}$$

• model B:

$$\mathbf{W}_2 = \frac{1}{2}\,(00011000)\,(10100000)\ , \qquad \mathbf{W}_3 = \frac{1}{3}\,(01-100000)\,(00200000)\ . \tag{B5}$$

One can easily verify our choices satisfy the modular-invariance requirements, eqs. A10 and A11. The $\mathbf{W}_3$ Wilson line of models A1/B and A2 breaks the gauge group in the observable sector to SO_{10} and $\mathrm{SO}_{10}\times\mathrm{SU}_2$, respectively, and the $\mathbf{W}_2$ breaks them further down to the PS gauge group, in all three models.

The remaining unbroken gauge groups are $\mathrm{SU}_{4\mathrm{C}}\times\mathrm{SU}_{2\mathrm{L}}\times\mathrm{SU}_{2\mathrm{R}}\times\mathrm{SO}_{10}{}'\times\mathrm{SU}_2{}'\times(\mathrm{U}_1)^5$ for models A1/A2 and $\mathrm{SU}_{4\mathrm{C}}\times\mathrm{SU}_{2\mathrm{L}}\times\mathrm{SU}_{2\mathrm{R}}\times\mathrm{SO}_{10}{}'\times(\mathrm{U}_1)^6$ for model B. The untwisted- and twisted-sector matter provide the following irreducible representations of the PS gauge group (modulo some singlets),

• model A1:

$$\begin{aligned}
&U_1: \quad (\mathbf{4},\mathbf{2},\mathbf{1})\,,\ U_2: (\mathbf{1},\mathbf{2},\mathbf{2})\,,\ U_3: (\mathbf{4},\mathbf{1},\mathbf{2}) + (\overline{\mathbf{4}},\mathbf{1},\mathbf{2})\,,\\
&T_1: \quad 2(\mathbf{4},\mathbf{2},\mathbf{1}) + 2(\overline{\mathbf{4}},\mathbf{1},\mathbf{2}) + 4(\mathbf{4},\mathbf{1},\mathbf{1}) + 4(\overline{\mathbf{4}},\mathbf{1},\mathbf{1}) + 8(\mathbf{1},\mathbf{2},\mathbf{1}) + 8(\mathbf{1},\mathbf{1},\mathbf{2}) + 2(\mathbf{1},\mathbf{1},\mathbf{2};\mathbf{1},\mathbf{2})\,,\\
&T_2: \quad 2(\overline{\mathbf{4}},\mathbf{1},\mathbf{2}) + (\mathbf{6},\mathbf{1},\mathbf{1})\,, \qquad T_3: 6(\mathbf{6},\mathbf{1},\mathbf{1}) + 6(\mathbf{1},\mathbf{2},\mathbf{2})\,, \qquad T_4: (\mathbf{4},\mathbf{1},\mathbf{2}) + 2(\mathbf{6},\mathbf{1},\mathbf{1})\,,
\end{aligned} \tag{B6}$$

• model A2:

$$\begin{aligned}
U_1:&\quad (\mathbf{4},\mathbf{2},\mathbf{1})\,,\\
T_1:&\quad 2(\mathbf{4},\mathbf{2},\mathbf{1}) + 2(\overline{\mathbf{4}},\mathbf{1},\mathbf{2}) + 4(\mathbf{4},\mathbf{1},\mathbf{1}) + 4(\overline{\mathbf{4}},\mathbf{1},\mathbf{1}) + 8(\mathbf{1},\mathbf{2},\mathbf{1}) + 6(\mathbf{1},\mathbf{1},\mathbf{2})\\
&+\ 2(\mathbf{1},\mathbf{2},\mathbf{1};\mathbf{1},\mathbf{2}) + 2(\mathbf{1},\mathbf{1},\mathbf{2};\mathbf{1},\mathbf{2})\,,\\
T_2:&\quad 2(\overline{\mathbf{4}},\mathbf{1},\mathbf{2}) + (\mathbf{6},\mathbf{1},\mathbf{1}) + (\mathbf{1},\mathbf{2},\mathbf{2})\,, \qquad T_3: 2(\mathbf{4},\mathbf{1},\mathbf{1}) + 2(\overline{\mathbf{4}},\mathbf{1},\mathbf{1}) + 6(\mathbf{1},\mathbf{1},\mathbf{2})\,,\\
T_4:&\quad (\mathbf{4},\mathbf{1},\mathbf{2}) + 2(\mathbf{6},\mathbf{1},\mathbf{1}) + 2(\mathbf{1},\mathbf{2},\mathbf{2})\,,
\end{aligned} \tag{B7}$$

• model B:

$$
\begin{aligned}
U_1:&\quad (\mathbf{4},\mathbf{2},\mathbf{1})\,,\ U_2:(\mathbf{6},\mathbf{1},\mathbf{1})\,,\ U_3:(\overline{\mathbf{4}},\mathbf{2},\mathbf{1})+(\mathbf{4},\mathbf{2},\mathbf{1})\,,\\
T_1:&\quad 2(\mathbf{6},\mathbf{1},\mathbf{1})+2(\mathbf{1},\mathbf{2},\mathbf{2})+4(\mathbf{4},\mathbf{1},\mathbf{1})+4(\overline{\mathbf{4}},\mathbf{1},\mathbf{1})+10(\mathbf{1},\mathbf{2},\mathbf{1})+8(\mathbf{1},\mathbf{1},\mathbf{2})\,,\\
T_2:&\quad 2(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})+(\mathbf{1},\mathbf{2},\mathbf{2})\,,\qquad T_3:\ 4(\mathbf{4},\mathbf{2},\mathbf{1})+4(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})+2(\overline{\mathbf{4}},\mathbf{2},\mathbf{1})+2(\mathbf{4},\mathbf{1},\mathbf{2})+6(\mathbf{1},\mathbf{2},\mathbf{1})\,,\\
T_4:&\quad (\mathbf{4},\mathbf{1},\mathbf{2})+2(\mathbf{1},\mathbf{2},\mathbf{2})\,.
\end{aligned}
\tag{B8}
$$

In general, the T_{6-k}-sector is the CPT conjugate of the T_k-sector, therefore T_5 does not give rise to additional states. However we need to keep those states from the T_4-sector whose CPT conjugates in the T_2-sector have positive helicities.

We see there are indeed three chiral families ($(\mathbf{4},\mathbf{2},\mathbf{1})+(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})$ under the PS group), two from the T_1 (T_3) twisted sector and one from the untwisted and the $T_{2,4}$ twisted sectors in model A1/A2 (B), modulo some $(\mathbf{4},\mathbf{1},\mathbf{2})+(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})$ (and $(\mathbf{4},\mathbf{2},\mathbf{1})+(\overline{\mathbf{4}},\mathbf{2},\mathbf{1})$ in model B) vector-like pairs. Each $(\mathbf{4},\mathbf{2},\mathbf{1})+(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})$ family encompasses one complete family of the SM quarks and leptons (plus the right-handed neutrino), since they decompose under the SM gauge group, $\mathrm{SU_{3C}}\times\mathrm{SU_{2L}}\times\mathrm{U_{1Y}}$, as follows: $(\mathbf{4},\mathbf{2},\mathbf{1})=(\mathbf{3},\mathbf{2})_{1/6}+(\mathbf{1},\mathbf{2})_{-1/2}$, $(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})=(\overline{\mathbf{3}},\mathbf{1})_{1/3}+(\overline{\mathbf{3}},\mathbf{1})_{-2/3}+(\mathbf{1},\mathbf{1})_1+(\mathbf{1},\mathbf{1})_0$. The complete matter spectra for these models is listed in appendix B 2, where the notation for the twisted-sector states is also explained. Three-family PS models have been previously constructed in heterotic string models using the free-fermionic method [14], but to our knowledge, has not been realized with abelian orbifolds.

Like all the three-family orbifold models in the literature [9, 10, 11], our models suffer from the embarrassment-of-riches syndrome, i.e., they contain many candidates for SM exotic particles, $(\mathbf{4},\mathbf{1},\mathbf{1})$, $(\overline{\mathbf{4}},\mathbf{1},\mathbf{1})$ and $(\mathbf{1},\mathbf{2},\mathbf{1})$, $(\mathbf{1},\mathbf{1},\mathbf{2})$. (These states decompose under the SM gauge group as follows: $(\mathbf{4},\mathbf{1},\mathbf{1})=(\mathbf{3},\mathbf{1})_{1/6}+(\mathbf{1},\mathbf{1})_{-1/2}$, $(\overline{\mathbf{4}},\mathbf{1},\mathbf{1})=(\overline{\mathbf{3}},\mathbf{1})_{-1/6}+(\mathbf{1},\mathbf{1})_{1/2}$, $(\mathbf{1},\mathbf{2},\mathbf{1})=(\mathbf{1},\mathbf{2})_0$ and $(\mathbf{1},\mathbf{1},\mathbf{2})=(\mathbf{1},\mathbf{1})_{1/2}+(\mathbf{1},\mathbf{1})_{-1/2}$, and have exotic charge assignments.) Nevertheless, these exotic states are necessary for the quantum consistency of the models. Models A1/A2 (B) contain five (six) additional abelian symmetries (their charges for the matter fields are listed in appendix B 2), one of them is anomalous. This $\mathrm{U_{1A}}$ anomaly satisfies the following universal condition,

$$
\mathrm{Tr}T(R)\widetilde{Q}_A=\mathrm{Tr}\widetilde{Q}_i^2\widetilde{Q}_A=\frac{1}{3}\mathrm{Tr}\widetilde{Q}_A^3=\frac{1}{24}\mathrm{Tr}\widetilde{Q}_A=\begin{cases}-\frac{1}{2\sqrt{3}}\,, & \text{Model A1}\,,\\ 1\,, & \text{Model A2}\,,\\ \frac{4}{3}\,, & \text{Model B}\,,\end{cases}
\tag{B9}
$$

where $\widetilde{Q}_{i,A}=Q_{i,A}/k_{i,A}^{1/2}$ with k_i,k_A are the normalization factors ("levels") of the corresponding abelian groups, and $2T(R)$ is the index of the representation R under a specific non-abelian factor.

This condition guarantees the anomalies are cancelled by the generalized Green-Schwarz mechanism [42] in 4d [43]. In fact, considerations of these anomalies provide useful consistency checks for our computer program. One can also check that all the 4d chiral anomalies for the non-abelian gauge groups and global anomalies for the SU_2 groups vanish.

We note that our models differ from the three-family SM-like models constructed from $\mathbb{Z}_3$ orbifolds in several respects. For example, in the simplest model of the first paper in ref. [9], the left-handed quarks arise solely from the untwisted sectors, one family each from the three untwisted sectors. The right-handed quarks and left-handed leptons arise from the twisted sectors, and the multiplicities of three come from the three equivalent fixed points in one of the complex planes. As such, the three families are completely degenerate in the horizontal family space. Our models exhibit quite different spectral patterns. Two complete families come from the T_1/T_3 twisted sectors, and one family comes from the combination of the untwisted and T_2/T_4 twisted sectors. This type of pattern breaks the degeneracies among the three families, and may have better prospects for explaining the observed fermion mass hierarchy and CP phases. It will also be interesting to generalize our models to other non-prime-order and $\mathbb{Z}_N \times \mathbb{Z}_M$ orbifolds to see whether this feature persists.

The above three-family PS models have some advantages over string-based SM-like and GUT models which have been the foci of model buildings in the past. The SM-like models usually contain many extra abelian gauge symmetries, and in general it is difficult to reproduce the standard hypercharge normalization as in GUT theories. (For an extensive search of $\mathrm{U}_{1\mathrm{Y}}$ in the $\mathbb{Z}_3$ orbifold, see ref. [50].) This may upset the successful prediction of the weak-mixing angle in the context of supersymmetric unification theories (see, however, ref. [51] for alternative tri-unification models). In the PS model, on the other hand, the hypercharge normalization is standard, by construction. The hypercharge $\mathrm{U}_{1\mathrm{Y}}$ is a diagonal subgroup of the $\mathrm{U}_{1\mathrm{T}_3} \subset \mathrm{SU}_{2\mathrm{R}}$ and the $\mathrm{U}_{1(\mathrm{B-L})} \subset \mathrm{SU}_{4\mathrm{C}}$, i.e. $Q_Y = T_{3R} + \frac{1}{2}(B-L)$, hence $k_Y = 1 + \frac{1}{4} \times \frac{8}{3} = \frac{5}{3}$. Moreover, all our models contain SO_{10} spinor representations, which may be a welcoming ingredient for better low energy phenomenology [52]. (In contrast, the spinor representations are absent in SM-like string models and models based on the $\mathrm{Spin}(32)/\mathbb{Z}_2$ heterotic string and superstrings.) Finally, when compared with string GUT models based on SU_5 or other larger gauge symmetries [13], the PS models generically do not need large Higgs representations (such as those in the adjoint representation) to break the GUT gauge symmetries. These large Higgs representations require one to realize the current algebra of the gauge symmetries at higher Kač-Moody levels [53] and render the models more complicated.

2. Complete matter spectra

We list the complete matter content of the three-family PS models, including all the singlet fields, in tables II, III and IV.

Let us explain our notation for the twisted-sector states somewhat. For the sake of presentation, we take the six compactified dimensions to be a factorizable Lie algebra root lattice, $\mathrm{G}_2 \oplus \mathrm{SU}_3 \oplus \mathrm{SO}_4$ (see fig. 2), as in sect. III A.[17] The notation, although pertaining to the particular case, can be generalized to any other lattice.

The twisted sectors contain various matter states, arising from different fixed points. It is useful to introduce the following notation for the twisted-sector states (see appendix A, also ref. [48]),

$$T_{k(\gamma)_\alpha (n_3)(n_2 n_2')N}, \tag{B10}$$

where the first sub-index represents the k^{th} twisted sector, the next three quantum numbers, $(\gamma)_\alpha$, (n_3) and (n_2, n_2'), specify fixed points on the first, second and third complex planes, respectively. The last sub-index $N = (N_1 N_2 N_3)$ denotes an array of three integral oscillator numbers for the left movers. For example, (100) and $(\overline{1}00)$ represent creating one oscillator in the Z_1 and $\overline{Z}_1$ directions. We only need to keep the left-moving oscillator numbers since all the massless states in our models have zero numbers of right-moving oscillators.

The quantum numbers $(\gamma)_\alpha$, (n_3) and (n_2, n_2') label conjugacy classes of the fixed points. The second complex plane, i.e., the SU_3 lattice, is a $\mathbb{Z}_3$ orbifold, and has three fixed points, $\mathbf{f} = \mathbf{0}$, $\frac{1}{3}(2\mathbf{e}_3 + \mathbf{e}_4)$ and $\frac{1}{3}(\mathbf{e}_3 + 2\mathbf{e}_4)$, where $\mathbf{e}_3$, $\mathbf{e}_4$ are the basis of the SU_3 lattice. The conjugacy classes are labelled by equivalent shift vectors $a\mathbf{e}_3 + b\mathbf{e}_4$, where $a + b \equiv n_3 = 0, 1, 2 \bmod 3$ is the "winding number." We can introduce at most one degree-3 Wilson line, and specify these fixed points by $n_3 = 0, 1, 2$ for $k \neq 3$, and $n_3 = 0$ for $k = 3$ (since Z_2 is the fixed plane of the θ^3 twist). Similarly, the third complex plane, i.e., the SO_4 lattice, is a $\mathbb{Z}_2$ orbifold, and has four fixed points, $\mathbf{f} = \frac{1}{2}(n_2 \mathbf{e}_5 + n_2' \mathbf{e}_6)$, where $n_2, n_2' = 0, 1$ and $\mathbf{e}_5$, $\mathbf{e}_6$ are the basis vectors of the SO_4 lattice. The equivalent shift vectors for these fixed points are $-n_2 \mathbf{e}_5 - n_2' \mathbf{e}_6$. We can introduce at most two

[17] In fact, the matter spectra also depend on our choice of the compactifed lattice. The Lie algebra root lattice sits at special points of the six-torus modulus space with enhanced gauge symmetries. The spectra (if we restrict ourselves to the root lattice) should be regarded as truncated spectra of the complete model, where there are additional states neutral under the observable gauge groups. The use of non-simply-laced algebra G_2 may also be worrisome. But it may be constructed from a simply-laced algebra such as the SO_8 by appropriate outer-automorphism, at the expense of reducing the rank of the hidden-sector groups. The G_2 lattice can also be replaced by the $\mathrm{SU}_3{}^{[2]}$ lattice without changing any of our discussions on low energy phenomenology. The $\mathrm{SU}_3{}^{[2]}$ Coxeter element is generated by the Weyl reflections with respect to $\mathbf{e}_1$, s_1, and with respect to $\mathbf{e}_2 - \mathbf{e}_1$, $s_{2\leftrightarrow 1}$, $C_{\mathrm{SU}_3^{[2]}} = s_1 s_{2\leftrightarrow 1} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$, where $\mathbf{e}_1$ and $\mathbf{e}_2$ are the simple roots of the SU_3.

TABLE II: Matter spectrum of model A1. The levels of the U_1 groups are 24, 24, 72, 36 and 48.

Sectors	$\mathrm{PS}\times\mathrm{SO}'_{10}\times\mathrm{SU}'_2$	Q_1	Q_2	Q_3	Q_4	Q_A	Labels	Sectors	$\mathrm{PS}\times\mathrm{SO}'_{10}\times\mathrm{SU}'_2$	Q_1	Q_2	Q_3	Q_4	Q_A	Labels
U_1	$(\mathbf{4},\mathbf{2},\mathbf{1},\mathbf{1},\mathbf{1})$	1	1	0	3	−2	f_3	$T_{2(-1)(1)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	−2	−4	−2	0	S_{14}
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{10},\mathbf{2})$	0	0	0	1	0	B'	$T_{2(-1)(2)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	4	−2	4	S_{15}
U_2	$(\mathbf{1},\mathbf{2},\mathbf{2},\mathbf{1},\mathbf{1})$	2	−1	0	−2	2	h_1		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	−2	−2	S_{16}
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	2	0	S_1		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{10},\mathbf{1})$	0	0	−2	2	−2	A'_1
U_3	$(\mathbf{4},\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	3	0	0	1	0	$\overline{\chi}^c_1$	$T_{2(1)(0)(00)(000)}$	$(\overline{\mathbf{4}},\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	1	0	0	1	0	$\chi^c_{1,2}$
	$(\overline{\mathbf{4}},\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	−3	0	0	−1	0	f^c_3	$T_{2(1)(1)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{2})$	0	−2	−2	−5	0	D'_5
$T_{1(1)(0)(0n')(000)}$	$(\mathbf{4},\mathbf{2},\mathbf{1},\mathbf{1},\mathbf{1})$	−1	0	0	0	0	$f_{1,2}$	$T_{2(1)(2)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{2})$	0	0	2	−1	4	D'_6
	$(\overline{\mathbf{4}},\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	−1	0	0	0	0	$f^c_{1,2}$		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{16},\mathbf{1})$	0	0	−1	0	−2	F'
$T_{1(1)(1)(0n')(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{2})$	−2	0	−2	0	−2	D'_1	$T_{2(-1)(0)(00)(0\overline{1}0)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	2	4	4	2	S_{17}
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{2})$	2	1	2	2	2	D'_2	$T_{2(-1)(2)(00)(0\overline{1}0)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	0	−2	S_{18}
$T_{1(1)(2)(0n')(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{2})$	−2	1	2	0	0	D'_3	$T_{2(-1)(2)(00)(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{2})$	0	0	−2	1	−2	D'_7
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{2})$	2	−1	−2	−2	−2	D'_4	$T_{2(1)(0)(00)(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	2	4	4	2	S_{19}
$T_{1(1)(0)(1n')(000)}$	$(\mathbf{4},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	−1	0	2	−1	3	q_1	$T_{2(1)(2)(00)(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	0	−2	S_{20}
	$(\mathbf{4},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	−1	0	−2	1	−3	q_2	$T_{2(1)(2)(00)(0\overline{1}0)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{2})$	0	0	−2	1	−2	D'_8
	$(\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	2	0	3	D^ℓ_1	$T_{3(\omega)(0)(0n')(000)}$	$(\mathbf{6},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	1	0	C_2
	$(\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	−2	2	−3	D^ℓ_2		$(\mathbf{1},\mathbf{2},\mathbf{2},\mathbf{1},\mathbf{1})$	0	0	0	1	0	h_2
$T_{1(1)(1)(1n')(000)}$	$(\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1},\mathbf{1})$	−2	0	2	−2	1	D^ℓ_3		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	−1	0	−3	2	S_{21}
	$(\mathbf{1},\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	0	−1	−2	−2	−3	D^r_1	$T_{3(\omega^2)(0)(0n')(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	−2	1	0	1	−2	S_{22}
	$(\overline{\mathbf{4}},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	1	0	2	−1	1	$\overline{q}_1$		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	−1	0	−1	2	S_{23}
$T_{1(1)(2)(1n')(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{2})$	0	0	0	−1	−1	Δ	$T_{3(1)(0)(0n')(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	−2	1	0	3	−2	S_{24}
$T_{1(1)(0)(0n')(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	−2	−1	0	−3	2	S_2		$(\mathbf{6},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	−1	0	C_3
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	−2	−1	−4	−1	−4	S_3		$(\mathbf{1},\mathbf{2},\mathbf{2},\mathbf{1},\mathbf{1})$	0	0	0	−1	0	h_3
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	−2	2	4	3	2	S_4	$T_{4(-1)(0)(00)(000)}$	$(\mathbf{4},\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	−1	0	0	−1	0	$\overline{\chi}^c_2$
$T_{1(1)(1)(0n')(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	−2	0	0	−1	−2	S_5	$T_{4(-1)(1)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{2})$	0	2	2	5	0	D'_9
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	1	4	1	2	S_6	$T_{4(-1)(2)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{2})$	0	0	−2	1	−4	D'_{10}
$T_{1(1)(2)(0n')(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	−2	1	0	3	0	S_7		$(\mathbf{1},\mathbf{1},\mathbf{1},\overline{\mathbf{16}},\mathbf{1})$	0	0	1	0	2	$\overline{F}'$
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	−1	−4	1	−2	S_8	$T_{4(1)(0)(00)(000)}$	$(\mathbf{6},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	0	0	0	C_4
$T_{1(1)(0)(1n')(100)}$	$(\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1},\mathbf{1})$	−2	−1	−2	−2	−1	D^ℓ_4		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	−2	−4	−2	−2	S_{25}
	$(\overline{\mathbf{4}},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	1	−1	−2	−1	−1	$\overline{q}_2$		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	−4	0	0	−2	0	S_{26}
$T_{1(1)(2)(1n')(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	0	0	−2	2	−1	D^r_2	$T_{4(1)(1)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	2	4	2	0	S_{27}
$T_{1(1)(0)(1n')(010)}$	$(\mathbf{1},\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	0	1	2	2	1	D^r_3	$T_{4(1)(2)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	−4	2	−4	S_{28}
$T_{1(1)(0)(0n')(001)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	0	1	0	S_9		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	2	2	S_{29}
$T_{1(1)(0)(0n')(00\overline{1})}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	0	1	0	S'_9		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{10},\mathbf{1})$	0	0	2	−2	2	A'_2
$T_{1(1)(0)(1n')(200)}$	$(\mathbf{1},\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	0	1	2	2	1	D^r_4	$T_{4(-1)(0)(00)(\overline{1}00)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	−2	−4	−4	−2	S_{30}
$T_{1(1)(0)(0n')(110)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	0	1	0	S_{10}	$T_{4(-1)(2)(00)(\overline{1}00)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	0	2	S_{31}
$T_{1(1)(0)(0n')(300)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	0	1	0	S_{11}	$T_{4(-1)(2)(00)(010)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{2})$	0	0	2	−1	2	D'_{11}
$T_{2(-1)(0)(00)(000)}$	$(\mathbf{6},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	−2	0	0	0	0	C_1	$T_{4(1)(0)(00)(010)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	−2	−4	−4	−2	S_{32}
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	2	4	2	2	S_{12}	$T_{4(1)(2)(00)(010)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	0	2	S_{33}
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	4	0	0	2	0	S_{13}	$T_{4(1)(2)(00)(\overline{1}00)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{2})$	0	0	2	−1	2	D'_{12}

TABLE III: Matter spectrum of model A2. The levels of the U_1 groups are 16, 72, 32, 80 and 36.

Sectors	$PS \times SO_{10}' \times SU_2'$	Q_1	Q_2	Q_3	Q_4	Q_A
U_1	$(\mathbf{4,2,1,1,1})$	0	3	1	−5	0
	$(\mathbf{1,1,1,16,1})$	0	0	1	1	3
	$(\mathbf{1,1,1,1,2})$	0	0	−2	4	3
U_2	$(\mathbf{1,1,1,1,1})$	2	−6	−4	0	0
$T_{1(1)(0)(0n')(000)}$	$(\mathbf{4,2,1,1,1})$	0	−1	−1	1	1
	$(\overline{\mathbf{4}},\mathbf{1,2,1,1})$	0	−1	−1	1	1
$T_{1(1)(1)(0n')(000)}$	$(\mathbf{1,1,1,16,1})$	0	0	−1	−1	0
$T_{1(1)(2)(0n')(000)}$	$(\mathbf{1,1,1,1,2})$	−2	4	0	0	2
	$(\mathbf{1,1,1,1,2})$	2	−2	0	2	2
$T_{1(1)(0)(1n')(000)}$	$(\mathbf{4,1,1,1,1})$	−1	2	−1	−3	1
	$(\mathbf{4,1,1,1,1})$	1	−4	−1	5	1
	$(\mathbf{1,2,1,1,1})$	−1	5	2	−4	1
	$(\mathbf{1,2,1,1,1})$	1	−1	2	4	1
$T_{1(1)(1)(1n')(000)}$	$(\mathbf{1,2,1,1,2})$	1	−3	0	2	0
	$(\mathbf{1,1,2,1,2})$	−1	3	2	−2	0
$T_{1(1)(2)(1n')(000)}$	$(\mathbf{1,1,2,1,1})$	−1	1	0	−4	−1
	$(\mathbf{1,2,1,1,1})$	−1	1	2	0	−1
	$(\overline{\mathbf{4}},\mathbf{1,1,1,1})$	1	−2	1	−1	−1
$T_{1(1)(0)(0n')(100)}$	$(\mathbf{1,1,1,1,1})$	−2	2	−2	−2	1
	$(\mathbf{1,1,1,1,1})$	0	−4	−2	6	1
	$(\mathbf{1,1,1,1,1})$	2	−4	−2	0	1
$T_{1(1)(1)(0n')(100)}$	$(\mathbf{1,1,1,1,1})$	0	0	−2	−2	−3
	$(\mathbf{1,1,1,1,1})$	0	0	−2	2	3
$T_{1(1)(2)(0n')(100)}$	$(\mathbf{1,1,1,1,1})$	−2	4	2	−4	−1
	$(\mathbf{1,1,1,1,1})$	2	−2	2	−2	−1
$T_{1(1)(0)(1n')(100)}$	$(\mathbf{1,2,1,1,1})$	−1	−1	−2	2	1
	$(\overline{\mathbf{4}},\mathbf{1,1,1,1})$	−1	2	1	1	1
$T_{1(1)(1)(0n')(010)}$	$(\mathbf{1,1,1,1,2})$	0	0	0	−2	0
$T_{1(1)(0)(1n')(010)}$	$(\mathbf{1,1,2,1,1})$	1	−1	0	0	1
$T_{1(1)(0)(0n')(001)}$	$(\mathbf{1,1,1,1,1})$	0	2	2	0	1
$T_{1(1)(0)(0n')(00\overline{1})}$	$(\mathbf{1,1,1,1,1})$	0	2	2	0	1
$T_{1(1)(1)(0n')(200)}$	$(\mathbf{1,1,1,1,2})$	0	0	0	−2	0
$T_{1(1)(0)(1n')(200)}$	$(\mathbf{1,1,2,1,1})$	1	−1	0	0	1
$T_{1(1)(0)(0n')(110)}$	$(\mathbf{1,1,1,1,1})$	0	2	2	0	1
$T_{1(1)(0)(0n')(300)}$	$(\mathbf{1,1,1,1,1})$	0	2	2	0	1
$T_{2(-1)(0)(00)(000)}$	$(\mathbf{6,1,1,1,1})$	0	−2	−2	2	2
	$(\mathbf{1,1,1,1,1})$	2	−2	0	−4	−4
	$(\mathbf{1,1,1,1,1})$	0	4	4	0	2
$T_{2(-1)(1)(00)(000)}$	$(\mathbf{1,1,1,10,1})$	0	0	−2	−2	0

Sectors	$PS \times SO_{10}' \times SU_2'$	Q_1	Q_2	Q_3	Q_4	Q_A
$T_{2(-1)(1)(00)(000)}$	$(\mathbf{1,1,1,1,1})$	2	−6	0	4	0
	$(\mathbf{1,2,2,1,1})$	0	0	2	0	0
	$(\mathbf{1,1,1,1,1})$	−2	6	4	−4	0
$T_{2(-1)(2)(00)(000)}$	$(\mathbf{1,1,1,1,1})$	0	−4	−4	4	−2
$T_{2(1)(0)(00)(000)}$	$(\overline{\mathbf{4}},\mathbf{1,2,1,1})$	0	1	1	1	2
$T_{2(1)(1)(00)(000)}$	$(\mathbf{1,1,1,1,2})$	0	0	−2	−4	−3
	$(\mathbf{1,1,1,1,2})$	0	0	−2	0	3
$T_{2(1)(2)(00)(000)}$	$(\mathbf{1,1,1,1,2})$	0	−4	−2	4	1
$T_{2(-1)(0)(00)(0\overline{1}0)}$	$(\mathbf{1,1,1,1,1})$	2	−2	0	0	2
$T_{2(-1)(1)(00)(0\overline{1}0)}$	$(\mathbf{1,1,1,1,1})$	0	0	0	−4	0
$T_{2(1)(0)(00)(100)}$	$(\mathbf{1,1,1,1,1})$	2	−2	0	0	2
$T_{2(1)(1)(00)(100)}$	$(\mathbf{1,1,1,1,1})$	0	0	0	−4	0
$T_{3(\omega)(0)(0n')(000)}$	$(\mathbf{1,1,1,1,1})$	0	0	2	2	−3
$T_{3(\omega)(0)(1n')(000)}$	$(\mathbf{1,1,2,1,1})$	−1	3	0	−2	3
$T_{3(\omega^2)(0)(0n')(000)}$	$(\mathbf{1,1,1,1,1})$	−2	0	−2	0	−3
	$(\mathbf{1,1,1,1,1})$	2	−6	−2	2	−3
	$(\mathbf{1,1,1,1,1})$	−2	6	2	−2	3
	$(\mathbf{1,1,1,1,1})$	2	0	2	0	3
$T_{3(\omega^2)(0)(1n')(000)}$	$(\mathbf{4,1,1,1,1})$	−1	0	−1	3	3
	$(\overline{\mathbf{4}},\mathbf{1,1,1,1})$	1	0	1	−3	−3
$T_{3(1)(0)(0n')(000)}$	$(\mathbf{1,1,1,1,1})$	0	0	−2	−2	3
$T_{3(1)(0)(1n')(000)}$	$(\mathbf{1,1,2,1,1})$	1	−3	0	2	−3
$T_{4(-1)(0)(00)(000)}$	$(\mathbf{4,1,2,1,1})$	0	−1	−1	−1	−2
$T_{4(-1)(1)(00)(000)}$	$(\mathbf{1,1,1,1,2})$	0	0	2	4	3
	$(\mathbf{1,1,1,1,2})$	0	0	2	0	−3
$T_{4(-1)(2)(00)(000)}$	$(\mathbf{1,1,1,1,2})$	0	4	2	−4	−1
$T_{4(1)(0)(00)(000)}$	$(\mathbf{6,1,1,1,1})$	0	2	2	−2	−2
	$(\mathbf{1,1,1,1,1})$	−2	2	0	4	4
	$(\mathbf{1,1,1,1,1})$	0	−4	−4	0	−2
$T_{4(1)(1)(00)(000)}$	$(\mathbf{1,1,1,10,1})$	0	0	2	2	0
	$(\mathbf{1,1,1,1,1})$	−2	6	0	−4	0
	$(\mathbf{1,2,2,1,1})$	0	0	−2	0	0
	$(\mathbf{1,1,1,1,1})$	2	−6	−4	4	0
$T_{4(1)(2)(00)(000)}$	$(\mathbf{1,1,1,1,1})$	0	4	4	−4	2
$T_{4(-1)(0)(00)(\overline{1}00)}$	$(\mathbf{1,1,1,1,1})$	−2	2	0	0	−2
$T_{4(-1)(1)(00)(\overline{1}00)}$	$(\mathbf{1,1,1,1,1})$	0	0	0	4	0
$T_{4(1)(0)(00)(010)}$	$(\mathbf{1,1,1,1,1})$	−2	2	0	0	−2
$T_{4(1)(1)(00)(010)}$	$(\mathbf{1,1,1,1,1})$	0	0	0	4	0

TABLE IV: Matter spectrum of model B. The levels of the U_1 groups are 24, 88, 40, 216, 88 and 576.

Sectors	$PS \times SO_{10}{}'$	Q_1	Q_2	Q_3	Q_4	Q_5	Q_A	Sectors	$PS \times SO_{10}{}'$	Q_1	Q_2	Q_3	Q_4	Q_5	Q_A
U_1	$(\mathbf{4},\mathbf{2},\mathbf{1},\mathbf{1})$	-1	0	1	0	0	12	$T_{2(1)(0)(00)(000)}$	$(\overline{\mathbf{4}},\mathbf{1},\mathbf{2},\mathbf{1})$	-1	0	-1	-4	0	-1
U_2	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	4	0	0	0	0	-12		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	-4	4	8	0	5
	$(\mathbf{6},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	0	-2	0	0	-12		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	4	0	8	0	11
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{10})$	0	-2	0	6	2	6	$T_{2(1)(1)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{10})$	0	2	0	-2	2	5
U_3	$(\mathbf{4},\mathbf{2},\mathbf{1},\mathbf{1})$	3	0	1	0	0	0		$(\mathbf{1},\mathbf{1},\mathbf{1},\overline{\mathbf{16}})$	0	1	0	1	-1	8
	$(\overline{\mathbf{4}},\mathbf{2},\mathbf{1},\mathbf{1})$	-3	0	-1	0	0	0		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	4	-4	11
$T_{1(1)(0)(0n')(000)}$	$(\mathbf{6},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	4	0	-5	$T_{2(1)(2)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{10})$	0	-2	0	-6	-2	5
	$(\mathbf{1},\mathbf{2},\mathbf{2},\mathbf{1})$	0	0	0	4	0	-5		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	4	12	4	5
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	0	-8	0	4	$T_{2(-1)(1)(00)(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	4	4	11
$T_{1(1)(1)(0n')(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	-2	-4	0	13	$T_{2(-1)(2)(00)(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	-4	0	0	0	11
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	4	0	-4	-4	4	$T_{2(1)(1)(00)(0\overline{1}0)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	4	4	11
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	4	2	8	4	7	$T_{2(1)(2)(00)(0\overline{1}0)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	-4	0	0	0	11
$T_{1(1)(2)(0n')(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	-4	2	0	4	7	$T_{3(\omega)(0)(0n')(000)}$	$(\overline{\mathbf{4}},\mathbf{2},\mathbf{1},\mathbf{1})$	-1	0	-1	0	0	-6
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	4	0	0	-2		$(\mathbf{4},\mathbf{1},\mathbf{2},\mathbf{1})$	-1	0	-1	0	0	-6
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	-4	2	0	-4	7	$T_{3(\omega)(0)(1n')(000)}$	$(\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	0	-2	4	6	2	-6
$T_{1(1)(0)(1n')(000)}$	$(\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	0	2	0	-2	2	-11	$T_{3(1)(0)(0n')(000)}$	$(\mathbf{4},\mathbf{2},\mathbf{1},\mathbf{1})$	1	0	1	0	0	6
	$(\mathbf{4},\mathbf{1},\mathbf{1},\mathbf{1})$	-1	-2	-1	-2	-2	10		$(\overline{\mathbf{4}},\mathbf{1},\mathbf{2},\mathbf{1})$	1	0	1	0	0	6
	$(\mathbf{1},\mathbf{1},\mathbf{2},\mathbf{1})$	2	-2	0	-2	-2	10	$T_{3(1)(0)(0n')(000)}$	$(\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	0	2	-4	-6	-2	6
$T_{1(1)(1)(1n')(000)}$	$(\mathbf{4},\mathbf{1},\mathbf{1},\mathbf{1})$	-1	2	-1	2	2	10	$T_{3(\omega)(0)(0n')(001)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	0	0	0	0	6
	$(\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	0	-2	2	2	2	13	$T_{3(\omega)(0)(0n')(\overline{1}00)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	0	0	0	-6
	$(\mathbf{1},\mathbf{1},\mathbf{2},\mathbf{1})$	2	2	0	2	2	10	$T_{3(\omega)(0)(0n')(00\overline{1})}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	0	0	0	-6
$T_{1(1)(2)(1n')(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{2},\mathbf{1})$	-2	2	-2	-6	-2	-2	$T_{3(\omega^2)(0)(0n')(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	0	0	0	-6
	$(\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	0	-2	0	-6	-2	-11	$T_{3(\omega^2)(0)(0n')(\overline{1}00)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	0	0	0	0	6
	$(\overline{\mathbf{4}},\mathbf{1},\mathbf{1},\mathbf{1})$	1	2	-1	-6	-2	-2	$T_{3(1)(0)(0n')(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	0	0	0	-6
$T_{1(1)(0)(0n')(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	-4	2	4	0	-2	$T_{3(1)(0)(0n')(001)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	0	0	0	0	6
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	4	-2	4	0	4	$T_{3(1)(0)(0n')(00\overline{1})}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	0	0	0	0	6
$T_{1(1)(1)(0n')(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	-4	-2	-4	0	4	$T_{4(-1)(0)(00)(000)}$	$(\mathbf{4},\mathbf{1},\mathbf{2},\mathbf{1})$	1	0	1	4	0	1
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	0	0	-4	-4	-5		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	4	-4	-8	0	-5
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	0	2	8	4	-2		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	-4	0	-8	0	-11
$T_{1(1)(2)(0n')(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	0	-2	0	4	4	$T_{4(-1)(1)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{10})$	0	-2	0	2	-2	-5
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	4	0	0	0	-5		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{16})$	0	-1	0	-1	1	-8
	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	-2	0	-2	0	-4	4		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	-4	4	-11
$T_{1(1)(0)(1n')(100)}$	$(\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	0	-2	-2	-2	-2	-2	$T_{4(-1)(2)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{10})$	0	2	0	6	2	-5
$T_{1(1)(1)(1n')(100)}$	$(\mathbf{1},\mathbf{1},\mathbf{2},\mathbf{1})$	-2	-2	0	2	2	1		$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	-4	-12	-4	-5
	$(\mathbf{1},\mathbf{2},\mathbf{1},\mathbf{1})$	0	2	-2	2	2	-2	$T_{4(1)(0)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	-8	0	10
	$(\overline{\mathbf{4}},\mathbf{1},\mathbf{1},\mathbf{1})$	1	-2	1	2	2	1		$(\mathbf{1},\mathbf{2},\mathbf{2},\mathbf{1})$	-2	0	0	4	0	1
$T_{1(1)(0)(0n')(200)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	2	4	0	7	$T_{4(1)(1)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	4	0	-4	-4	10
$T_{1(1)(0)(0n')(010)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	2	0	2	4	0	7	$T_{4(1)(2)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	4	0	0	4
$T_{2(-1)(0)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	8	0	-10	$T_{4(-1)(1)(00)(010)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	-4	-4	-11
	$(\mathbf{1},\mathbf{2},\mathbf{2},\mathbf{1})$	2	0	0	-4	0	-1	$T_{4(-1)(2)(00)(010)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	4	0	0	0	-11
$T_{2(-1)(1)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	-4	0	4	4	-10	$T_{4(1)(1)(00)(\overline{1}00)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	0	-4	-4	-11
$T_{2(-1)(2)(00)(000)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	0	-4	0	0	-4	$T_{4(1)(2)(00)(\overline{1}00)}$	$(\mathbf{1},\mathbf{1},\mathbf{1},\mathbf{1})$	0	4	0	0	0	-11

independent degree-2 Wilson lines, and specify these fixed points by a pair of winding numbers, $n_2, n_2' = 0, 1$ for $k = 1, 3, 5$, and $n_2 = n_2' = 0$ for $k = 2, 4$ (since Z_3 is the fixed plane of the $\theta^{2,4}$ twists). In the present models, we introduce only one non-vanishing $\mathbb{Z}_2$ Wilson line, so the degree-2 degeneracy associated with n_2' is not resolved.

For the first complex plane (i.e., the G_2 lattice), there is only one fixed point in the T_1 twisted sector, $\mathbf{f} = \mathbf{0}$. The corresponding state will be denoted by $(\gamma) = (1)$. In the T_2 and T_4 twisted sectors, there are three classes of fixed points, $\mathbf{f}_a = \frac{1}{3}a\mathbf{e}_1$ with $a = 0, 1, 2$. The last two, $\mathbf{f}_{1,2}$, transform into each other under the orbifold twist θ, while the first, $\mathbf{f}_0$, is invariant. For states associated with the non-invariant fixed points, we can take linear combinations such that they have definite θ-eigenvalues, $\gamma = \pm 1$. We label the two eigen-states with $\gamma = 1$ by $|\mathbf{f}_0\rangle \equiv (1)_1$, $\frac{1}{\sqrt{2}}(|\mathbf{f}_1\rangle + |\mathbf{f}_2\rangle) \equiv (1)_2$ and the one with $\gamma = -1$ by $\frac{1}{\sqrt{2}}(|\mathbf{f}_1\rangle - |\mathbf{f}_2\rangle) \equiv (-1)$. Similarly, there are four classes of fixed points in the T_3 twisted sector, $\mathbf{f}_{ab} = \frac{1}{2}(a\mathbf{e}_1 + b\mathbf{e}_2)$, with $a, b = 0, 1$. Three of them, $\mathbf{f}_{01}$, $\mathbf{f}_{10}$ and $\mathbf{f}_{11}$, transform into each other under the orbifold twist θ, while the other, $\mathbf{f}_{00}$, is invariant. After taking linear combinations, we have two states with $\gamma = 1$, $|\mathbf{f}_{00}\rangle$ and $\frac{1}{\sqrt{3}}(|\mathbf{f}_{10}\rangle + |\mathbf{f}_{01}\rangle + |\mathbf{f}_{11}\rangle)$, denoted by $(1)_{1,2}$, and two states $\frac{1}{\sqrt{3}}(|\mathbf{f}_{10}\rangle + \gamma|\mathbf{f}_{01}\rangle + \gamma^2|\mathbf{f}_{11}\rangle)$ with eigenvalues $\gamma = \omega, \omega^2$, where $\omega = \mathrm{e}^{\mathrm{i}2\pi/3}$, denoted by (ω) and (ω^2).

Multiplicities of the twisted-sector states are computed from the generalized GSO projector of eq. A14. They equal the number of degenerate fixed points with quantum numbers k, γ, α, n_3, n_2, n_2' and N. With two Wilson lines, one degree-2 and one degree-3, the T_1-sector states have multiplicities two, due to the unresolved degree-2 degeneracies in n_2'. The T_2 and T_4-sector states have multiplicities two and one, for states with eigenvalues $\gamma = 1$ and $\gamma = -1$, since they have two and one associated fixed points, $(1)_{1,2}$ and (-1). Finally, the T_3-sector states have multiplicities four and two, for states with eigenvalues $\gamma = 1$ and $\gamma = \omega, \omega^2$, since the associated fixed points are $(1)_{1,2}$ and $(\omega), (\omega^2)$ and there are additional degeneracies in $n_2' = 0, 1$.

APPENDIX C: YUKAWA COUPLINGS IN THE $\mathbb{Z}_6$ MODELS

1. Selection rules

In this subsection, we give the complete set of string selection rules in the $\mathbb{Z}_6$ orbifold model. Although these rules are well known, they are not always stated correctly in the previous literature.

In heterotic string models, physical states corresponding to space-time bosons can be written

in terms of vertex operators in the (-1)-ghost picture, they are

$$V_{-1}^{(\ell)} = \mathrm{e}^{-\phi} \prod_{i=1}^{3} (\partial Z_i)^{N_i^{(\ell)}} (\partial \overline{Z}_i)^{\overline{N}_i^{(\ell)}} \mathrm{e}^{-2\mathrm{i}\mathbf{r}^{(\ell)}\cdot\mathbf{H}} \mathrm{e}^{2\mathrm{i}\mathbf{P}^{(\ell)}\cdot\mathbf{X}} \sigma_f^{(\ell)}, \tag{C1}$$

where ϕ and $(\partial Z_i)^{N_i^{(\ell)}}$, $(\partial \overline{Z}_i)^{\overline{N}_i^{(\ell)}}$ denote the superconformal ghost and $N_i^{(\ell)}$ and $\overline{N}_i^{(\ell)}$ left-moving oscillators in the Z_i and $\overline{Z}_i$ directions. $\mathbf{P}^{(\ell)}$ and $\mathbf{r}^{(\ell)}$ are the $\mathrm{E}_8 \times \mathrm{E}_8$ and SO_8 shifted momenta ($\mathbf{X}$ and $\mathbf{H}$ are bosonic coordinates parameterizing the respective maximal tori); the latter is commonly known as the H-momentum. More explicitly, in the $\mathbb{Z}_6$ model the internal SO_6 parts of the H-momenta for states in different sectors are:

$$\begin{aligned} &U_1 :\ (1,0,0), \qquad U_2 :\ (0,1,0), \qquad U_3 :\ (0,0,1), \\ &T_1 :\ (1,2,3)/6, \qquad T_2 :\ (2,4,0)/6, \qquad T_3 :\ (3,0,3)/6, \qquad T_4 :\ (4,2,0)/6. \end{aligned} \tag{C2}$$

Finally, $\sigma_f^{(\ell)}$ denotes the twisted field, creating twisted-sector vacua out of the untwisted ones; it is thus trivial for the untwisted sector states.

Equivalently, due to the world-sheet superconformal symmetry for the right-moving superstring, bosonic state vertex operators can be written in other ghost pictures by acting with a picture changing operator [54], $\supset \mathrm{e}^{\phi} \sum_{i=1}^{3} (\psi_i \overline{\partial Z}_i + \overline{\psi}_i \overline{\partial} Z_i) = \mathrm{e}^{\phi} \sum_{i=1}^{3} (\mathrm{e}^{2\mathrm{i}\mathbf{r}_v^i\cdot\mathbf{H}} \overline{\partial Z}_i + \mathrm{e}^{-2\mathrm{i}\mathbf{r}_v^i\cdot\mathbf{H}} \overline{\partial} Z_i)$, where $\mathbf{r}_v^1 = (1,0,0)$, $\mathbf{r}_v^2 = (0,1,0)$, $\mathbf{r}_v^3 = (0,0,1)$. Therefore in the 0-ghost picture the H momenta are reduced by $\mathbf{r}_v$'s and additional factors of $\overline{\partial} Z_i$ and $\overline{\partial Z}_i$ are introduced. However the unique combination (defined up to a modding with respect to the degree of orbifolding)

$$R_i^{(\ell)} = r_i^{(\ell)} - N_i^{(\ell)} + \overline{N}_i^{(\ell)} \tag{C3}$$

remains invariant under picture-changing operations.

For Yukawa couplings in string theory, one can use the standard conformal field theory technique [54, 55] to compute the corresponding n-point correlation function,

$$\langle V_{-1}^{(1)} V_{-1/2}^{(2)} V_{-1/2}^{(3)} V_0^{(4)} V_0^{(5)} \cdots V_0^{(n)} \rangle, \tag{C4}$$

where $V_{-1/2}$'s are the vertex operators for space-time fermions in the $(-1/2)$-ghost picture and all but one of the bosonic vertex operators have been brought into the 0-ghost picture. Note that the total ghost charge must cancel with a background ghost charge of 2.

For our purposes, however, we do not need the exact form of the above correlation function;[18] string selection rules suffice to tell us whether certain types of Yukawa couplings are non-trivial, i.e., not identically zero. These rules are provided by different parts of the vertex operator.

[18] The functional dependence of Yukawa couplings on moduli can be determined by a field theoretical method

The gauge part requires conservation of the $E_8 \times E_8$ momenta, $\sum_\ell \mathbf{P}^{(\ell)} = 0$, which is nothing but the gauge invariance condition for matter couplings. A similar rule also holds for the H-momenta, but we have to be more careful because these momenta are not invariant under the picture-changing operations. Instead of the H-momenta, the rules must be imposed on the well-defined R-charges of eq. C3, they are

$$\sum_\ell R_1^{(\ell)} = 1 \bmod 6, \qquad \sum_\ell R_2^{(\ell)} = 1 \bmod 3, \qquad \sum_\ell R_3^{(\ell)} = 1 \bmod 2. \tag{C5}$$

We note that they can be understood as conservations of discrete global R-charges in the field theory language [60]. The nature of R-symmetries can be seen from the spacetime supersymmetry charge $Q_{\rm SUSY} = \int \frac{dz}{2\pi i} e^{-\phi/2} S e^{2i\mathbf{r}_{\rm SUSY}\cdot\mathbf{H}}$ where S is the spin field for uncompactified dimensions, $\mathbf{r}_{\rm SUSY} = \frac{1}{2}(1,1,1)$, and the integration is over the world-sheet coordinates. Under the orbifold twist $Q_{\rm SUSY} \to e^{i\pi v_i} Q_{\rm SUSY}$ and the bosonic vertex operator $V^{(\ell)} \to e^{-2i\pi R_i^{(\ell)} v_i} V^{(\ell)}$. Eq. C5 is equivalent to the statement that the superpotential is invariant under this twist.

The part associated with the twisted fields leads to the so-called space group selection rule. Global monodromy requires that the product of space group elements associated with the fixed points contain the unit $(1, \mathbf{0})$, that is,

$$\sum_\ell k_\ell = 0 \bmod 6\,, \tag{C6}$$

$$\sum_\ell (1 - \theta_{(\ell)}^{k_\ell}) \mathbf{f}_\ell = \sum_\ell (1 - \theta_{(\ell)}^{k_\ell}) \mathbf{\Lambda}_{(\ell)}\,, \tag{C7}$$

where the indices ℓ sum over the twisted-sector states (in the k_ℓ-th sector), $\mathbf{f}_\ell$ denote the fixed points under these twists, and $\mathbf{\Lambda}_\ell$ are arbitrary lattice vectors of the compactified space. Eq. C6 is the point group selection rule for the $\mathbb{Z}_6$ models.

Eq. C7 implicitly depends on the compactified lattice. For our choice of the root lattice $G_2 \oplus SU_3 \oplus SO_4$, using the notation of appendix B 2, we find it is equivalent to [48]

$$\sum_\ell n_3^{(\ell)} = 0 \bmod 3\,, \qquad \sum_\ell (n_2^{(\ell)}, n_2'^{(\ell)}) = (0,0) \bmod 2\,, \tag{C8}$$

[56]. The low energy effective action of 4d strings enjoys a target space duality symmetry $PSL(2,\mathbb{Z})$ [57] (or its congruence subgroups in cases with discrete Wilson lines [58]), under which the matter fields transform according to their modular weights. For the U_k untwisted-sector states, these weights are $t_i = -1$ (if $i = k$) and 0 (if $i \neq k$) for the three Kähler moduli. For twisted-sector states, $t_i = -1 + r_i - N_i + \overline{N}_i$, where $r_i, N_i, \overline{N}_i$ are the H-momenta and oscillator numbers. The superpotential has a weight $(-1,-1,-1)$ (since the combination of Kähler potential and superpotential $\mathcal{K} + \log|\mathcal{W}|^2$ has weight $(0,0,0)$), so the modulus-dependent Yukawa couplings must have certain definite weights under the duality group and they are the modular forms. Physically, the modulus dependence accounts for the world-sheet instanton effects [59] and may be important for suppressing some Yukawa couplings.

for quantum numbers in the second and third complex planes. We note the above rules can be understood as discrete global symmetries ($\mathbb{Z}_3$ and $\mathbb{Z}_2$ respectively) in the field theory language.

The conditions in the first plane are more complex since the fixed points need to be reorganized in terms of the θ-eigenstates. Denoting the states by $(\gamma^{(\ell)})_{\alpha^{(\ell)}}$, in addition to the apparent requirement that the product of θ-eigenvalues

$$\prod_{\ell} \gamma^{(\ell)} = 1\,, \tag{C9}$$

there are further conditions arising from eq. C7. We list them as follows:

1. There is no additional selection rules if the couplings involve states in the T_1 sector – all gauge-invariant couplings consistent with eqs. C5, C6 and C8 are allowed. This follows from the fact that in the T_1 sector all G_2 root vectors are in the same conjugacy class.

2. For couplings not involving any T_1 twisted state, the selection rule is determined by a $\mathbb{Z}_3$ ($\mathbb{Z}_2 \times \mathbb{Z}_2$) discrete symmetry for the $T_{2,4}$ (T_3) states. The $T_{2,4}$-sector states are related to the $\mathbb{Z}_3$ classes by $(\gamma = 1)_1 = [0]$, $(\gamma = 1)_2 = [1] + [2]$ and $(\gamma = -1) = [1] - [2]$, and the T_3-sector states are related to the $\mathbb{Z}_2 \times \mathbb{Z}_2$ classes by $(\gamma = 1)_1 = [00]$, $(\gamma = 1)_2 = [10] + [01] + [11]$ and $(\gamma = \omega, \omega^2) = [10]+\gamma[01]+\gamma^2[11]$. The allowed couplings can then be worked out by applying the appropriate multiplication tables of the conjugacy classes, $[a][b] = [(a + b) \bmod 3]$ and $[a_1 b_1][a_2 b_2] = [(a_1 + a_2) \bmod 2, (b_1 + b_2) \bmod 2]$.

The above selection rules can be used to determine the non-trivial Yukawa couplings. For example, consider three-point couplings (which correspond to renormalizable terms in the superpotential). From the H-momentum conservation eq. C5 and the point group rule eq. C6 we find the following set of allowed couplings in the $\mathbb{Z}_6$ model (all of these states have zero oscillator numbers),

$$U_1 U_2 U_3,\;\; T_1 T_2 T_3,\;\; T_1 T_1 T_4,\;\; T_3 T_3 U_2,\;\; T_2 T_4 U_3\,. \tag{C10}$$

Taking into account the space group rules, we find the complete list of three-point couplings in model A1 involving fields with non-trivial representations under the $\mathrm{SU}_{4\mathrm{C}}$ factor of the PS group (we have used the field notation of table II),

$$\begin{gathered}
h_1\, f_3\, f_3^c\,, \quad \sum_{\alpha,\beta=0,1} (h_3)_{A\alpha}\, f_B\, \chi_\beta^c\,, \quad C_1\, \overline{\chi}_1^c\, \overline{\chi}_2^c\,, \quad \sum_{\alpha,\beta=0,1} (C_3)_{A\alpha}\, f_B^c\, \chi_\beta^c\,, \\
\sum_{\alpha+\beta=0 \bmod 2} (C_4)_\alpha f_3^c\, \chi_\beta^c\,, \quad \sum_{\alpha=0,1} (C_4)_\alpha\, f_A\, f_B\,, \quad \sum_{\alpha=0,1} (C_4)_\alpha\, f_A^c\, f_B^c\,, \\
\sum_{\alpha+\beta=0 \bmod 2} (S_{26})_\alpha\, \chi_\beta^c\, \overline{\chi}_1^c\,, \quad S_{13}\, \overline{\chi}_2^c\, f_3^c\,, \quad \sum_{\alpha=0,1} (C_4)_\alpha\, (q_1)_A\, (q_2)_B\,,
\end{gathered}$$

$$\sum_{\alpha=0,1}(S_{28})_\alpha\,(q_1)_A\,(\overline{q}_1)_B\,,\qquad \sum_{\alpha+\beta=0\,\mathrm{mod}\,2} S_1\,(C_3)_{A\alpha}(C_3)_{B\beta}\,, \tag{C11}$$

where the "family" indices $A, B = 1, 2$ satisfy a $\mathbb{Z}_2$ condition, $A + B = 0\,\mathrm{mod}\,2$.

2. Allowed Yukawa couplings in the A1 model

In this subsection, we list allowed Yukawa couplings in the A1 model, using the string selection rules of appendix C 1. These operators all have the form $\mathcal{O}S^n$, where $\mathcal{O}$ are composite singlet operators that involve non-trivial PS fields, and S are the singlet fields. (Of course, one can also consider operators involving other observable- and hidden-sector fields.) We work out some of the lowest order allowed operators. We shall also follow the field naming scheme in table II.

First notice certain structures exist for these Yukawa couplings. A composite operator of singlet fields is equivalent to the unit operator $\mathbf{1}$ if it is inert under all the global and local symmetries of appendix C 1. Thus the superpotential factorizes $\mathcal{W} = \{\mathbf{1}\}\hat{\mathcal{W}}$, with the first few terms of $\{\mathbf{1}\}$ given by

$$\begin{aligned}
\{\mathbf{1}\} &= \mathbf{1} + S_{19}S_{32} + S_9(S_4S_{32} + S_5S_{33} + S_{10}S_{26}) + S_{19}^2S_{32}^2 + \cdots \\
&= (\mathbf{1} - S_{19}S_{32})^{-1}[\mathbf{1} - S_9(S_4S_{32} + S_5S_{33} + S_{10}S_{26})]^{-1} \times (\cdots),
\end{aligned} \tag{C12}$$

where we have performed a "resummation" in the second line. Note that $\{\mathbf{1}\}$ can be absorbed into the Kähler potential by a transformation, $\mathcal{K} \to \mathcal{K} + \log|\{\mathbf{1}\}|^2$. As a result, the sufficient conditions for F-flatness and vanishing cosmological constant are $\partial\hat{\mathcal{W}} = \hat{\mathcal{W}} = 0$. In the following expressions, we shall omit the caret.

We list the relevant effective operators as follows, suppressing the A and α indices (defined in the last section) except for those of the matter fields and the χ^c, $\overline{\chi}^c$ Higgs fields.

A. Color triplet masses

- The $(\mathbf{6},\mathbf{1},\mathbf{1})(\mathbf{6},\mathbf{1},\mathbf{1})$ operators:

$$\begin{aligned}
\mathcal{W} &= (S_1 + S_{19}S_{25} + S_9\mathcal{S}_1^{(2)} + S_2S_{24}\mathcal{S}_8^{(2)})(C_3)^2 + (S_2S_{24} + \mathcal{S}_1^{(2)}\mathcal{S}_2^{(2)})(C_4)^2 \\
&+ (\mathcal{S}_1^{(2)} + S_2S_9S_{24} + S_1\mathcal{S}_2^{(2)} + S_{19}S_{25}\mathcal{S}_2^{(2)})\,C_3C_4 \\
&+ S_2S_9S_{22}C_2C_4 + S_2S_{22}\mathcal{S}_8^{(2)}C_2C_3 + S_{10}\mathcal{S}_1^{(3)}(C_1)^2 + S_{10}\mathcal{S}_4^{(2)}\,C_1C_4 \\
&+ [S_{10}\mathcal{S}_3^{(2)} + S_{13}\mathcal{S}_2^{(2)} + S_1S_{10}S_{12}S_{32} + S_9S_{10}\mathcal{S}_4^{(2)}]\,C_1C_3 + \cdots
\end{aligned} \tag{C13}$$

- The $(\mathbf{6},\mathbf{1},\mathbf{1})(\mathbf{4},\mathbf{1},\mathbf{2})(\mathbf{4},\mathbf{1},\mathbf{2})$ operators:

$$\mathcal{W} = C_1\overline{\chi}_1^c\overline{\chi}_2^c + (S_2S_{12}S_{24}S_{32} + \mathcal{S}_2^{(2)}\mathcal{S}_4^{(2)})C_1(\overline{\chi}_1^c)^2$$

$$
\begin{aligned}
&+\ S_{10}(S_{10}\mathcal{S}_3^{(2)}+S_{13}\mathcal{S}_2^{(2)})C_1(\overline{\chi}_2^c)^2+S_2S_9S_{22}S_{26}C_2(\overline{\chi}_1^c)^2\\
&+\ S_{26}(\mathcal{S}_1^{(2)}+S_2S_9S_{24}+S_1\mathcal{S}_2^{(2)})C_3(\overline{\chi}_1^c)^2+(S_1S_{10}+S_{10}S_{19}S_{25}+S_9S_{10}\mathcal{S}_1^{(2)}+S_1\mathcal{S}_2^{(3)})C_3(\overline{\chi}_2^c)^2\\
&+\ S_2S_{24}S_{26}C_4(\overline{\chi}_1^c)^2+(S_{10}\mathcal{S}_1^{(2)}+S_6S_7S_{14}S_{18}+S_1S_{10}\mathcal{S}_2^{(2)}+S_2\mathcal{S}_4^{(3)})C_4(\overline{\chi}_2^c)^2+\cdots,
\end{aligned}
\tag{C14}
$$

- The $(\mathbf{6},\mathbf{1},\mathbf{1})(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})$ operators:

$$
\begin{aligned}
\mathcal{W}\ =&\ C_3f_A^c\chi_\alpha^c+C_4(f_3^c\chi_\alpha^c+f_A^cf_B^c)\\
&+\ (S_9+S_1S_{10}S_{21}S_{22}+\mathcal{S}_2^{(2)}S_{11}^2)[C_4f_3^cf_A^c+C_3(f_3^c\chi_\alpha^c+f_A^cf_B^c)]\\
&+\ (\mathcal{S}_8^{(2)}+S_{10}(S_{30}\mathcal{S}_5^{(2)}+S_{13}\mathcal{S}_4^{(2)}))[C_3f_3^cf_A^c+C_4(f_3^c)^2]+(\mathcal{S}_2^{(2)}+S_{26}\mathcal{S}_2^{(3)})[C_3\chi_\alpha^c\chi_\beta^c+C_4f_A^c\chi_\alpha^c],\\
&+\ S_{10}S_{12}S_{32}C_1f_A^c\chi_\alpha^c+[S_9\mathcal{S}_8^{(2)}+S_{13}(S_{10}\mathcal{S}_3^{(2)}+S_{13}\mathcal{S}_2^{(2)})]C_3(f_3^c)^2\\
&+\ S_{10}S_{13}S_{21}S_{22}C_1f_3^cf_A^c+S_9S_{10}S_{12}S_{32}C_1(f_3^c\chi_\alpha^c+f_A^cf_B^c)+(\mathcal{S}_2^{(2)})^2C_4\chi_\alpha^c\chi_\beta^c,
\end{aligned}
\tag{C15}
$$

$$
\mathcal{W}^{(8)}\ \supset\ S_{12}(S_{11}S_{30}\mathcal{S}_4^{(2)}+S_{10}S_{32}\mathcal{S}_2^{(2)})C_1\chi_\alpha^c\chi_\beta^c+S_2S_{10}S_{21}S_{22}^2C_2\chi_\alpha^c\chi_\beta^c.
\tag{C16}
$$

- The $(\mathbf{6},\mathbf{1},\mathbf{1})(\mathbf{4},\mathbf{2},\mathbf{1})(\mathbf{4},\mathbf{2},\mathbf{1})$ operators:

$$
\begin{aligned}
\mathcal{W}\ =&\ S_2^2S_9S_{10}S_{12}S_{32}C_1f_3^2+S_2S_{21}^2\mathcal{S}_7^{(2)}C_2f_3^2+S_2^2S_9C_3f_3^2+S_2^2C_4f_3^2\\
&+\ S_2S_9S_{10}S_{12}S_{32}C_1f_3f_A+S_{21}^2\mathcal{S}_7^{(2)}C_2f_3f_A+(S_2S_9+S_{21}S_{23}\mathcal{S}_7^{(2)})C_3f_3f_A+S_2C_4f_3f_A\\
&+\ S_9S_{10}S_{12}S_{32}C_1f_Af_B+S_{10}(S_2S_{12}S_{13}S_{22}S_{32}+S_3S_{12}^2S_{23}S_{32}+S_6S_{14}S_{16}S_{22}S_{33})C_2f_Af_B\\
&+\ (S_9+S_{11}^2\mathcal{S}_2^{(2)}+S_1S_{10}S_{21}S_{22})C_3f_Af_B+C_4f_Af_B+\cdots,
\end{aligned}
\tag{C17}
$$

In the above equations we have used the following composite singlet operators of dimensions two and three,

$$
\begin{aligned}
&\mathcal{S}_1^{(2)}=S_4S_{25}+S_5S_{29},\qquad \mathcal{S}_2^{(2)}=S_4S_{32}+S_5S_{33}+S_{10}S_{26},\\
&\mathcal{S}_3^{(2)}=S_{12}S_{25}+S_{13}S_{26}+S_{17}S_{32}+S_{19}S_{30},\qquad \mathcal{S}_4^{(2)}=S_4S_{30}+S_5S_{31},\\
&\mathcal{S}_5^{(2)}=S_4S_{13}+S_{10}S_{17},\qquad \mathcal{S}_6^{(2)}=S_{21}S_{24}+S_{22}S_{23},\\
&\mathcal{S}_7^{(2)}=S_3S_4+S_5S_7,\qquad \mathcal{S}_8^{(2)}=S_9^2+S_{11}^2,
\end{aligned}
\tag{C18}
$$

$$
\begin{aligned}
&\mathcal{S}_1^{(3)}=S_9S_{21}S_{22}+S_{10}S_{12}S_{30},\qquad \mathcal{S}_2^{(3)}=S_{10}S_{19}S_{32}+S_{11}S_{17}S_{30},\\
&\mathcal{S}_3^{(3)}=S_{10}S_{20}S_{33}+S_{11}S_{18}S_{31},\qquad \mathcal{S}_4^{(3)}=S_3S_{13}S_{17}+S_9S_{10}S_{24}.
\end{aligned}
\tag{C19}
$$

B. Quark and lepton Yukawa couplings

Quark and lepton Yukawa couplings are operators of the $(\mathbf{4},\mathbf{2},\mathbf{1})(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})(\mathbf{1},\mathbf{2},\mathbf{2})$ type. They are given by

$$
\mathcal{O}_{(a,b)}=f_ah_1f_b^c,\qquad a,\,b=1,2,3.
\tag{C20}
$$

In addition, the PS symmetry breaking fields may enter the effective fermion mass operators at higher order. We thus consider the operators of the $(\mathbf{4,1,2})(\overline{\mathbf{4}}\mathbf{,1,2})$ type given by

$$\mathcal{O}_1 = \overline{\chi}_1^c \chi_\alpha^c, \qquad \mathcal{O}_2 = \overline{\chi}_2^c \chi_\alpha^c \tag{C21}$$

Using these operators, we search for all the allowed higher dimensional operators of the form, $\mathcal{O}_{(a,b)}(\mathcal{O}_i)^n S^{n'}$, i.e., $(\mathbf{4},\mathbf{2},\mathbf{1})(\overline{\mathbf{4}},\mathbf{1},\mathbf{2})(\mathbf{1},\mathbf{2},\mathbf{2})[(\mathbf{4},\mathbf{1},\mathbf{2})(\overline{\mathbf{4}},\mathbf{2},\mathbf{1})]^n S^{n'}$, with the smallest possible n, n' (for $n+n' \leq 4$ or $n' \leq 5$). The result is $(A, B = 1, 2)$

$$\begin{aligned} \mathcal{W} \;=\; & \mathcal{S}_2^{(2)} \mathcal{O}_{(3,A)} + (S_{10}^2 S_{12} S_{24} S_{30} + S_9 S_{10} S_{22} \mathcal{S}_6^{(2)} + S_{21} S_{22} \mathcal{S}_4^{(3)} + S_3 S_{12} \mathcal{O}_1 \mathcal{O}_2) \mathcal{O}_{(A,3)} \\ +\; & (S_{10} S_{22} \mathcal{S}_6^{(2)} + S_3 S_9 S_{12} \mathcal{O}_2) \mathcal{O}_{(A,B)}. \end{aligned} \tag{C22}$$

C. Neutrino masses

Consider the three PS breaking operators

$$\mathcal{O}_{ij} = \overline{\chi}_i^c \overline{\chi}_j^c, \qquad i, j = 1, 2. \tag{C23}$$

We then find higher dimension effective Majorana neutrino mass operators of the form (with minimal number of singlets)

$$\begin{aligned} \mathcal{W} \;=\; & (\mathcal{S}_8^{(2)} S_{26} + S_1 S_{21} S_{22}) f_3^c f_3^c \mathcal{O}_{11} + S_{13} f_3^c f_3^c \mathcal{O}_{12} + S_9 S_{10} \mathcal{S}_8^{(2)} f_3^c f_3^c \mathcal{O}_{22} + S_9 S_{26} f_3^c f_A^c \mathcal{O}_{11} \\ +\; & (S_{10} \mathcal{S}_3^{(2)} + S_{13} \mathcal{S}_2^{(2)}) f_3^c f_A^c \mathcal{O}_{12} + S_{26} f_A^c f_B^c \mathcal{O}_{11} + S_{10} \mathcal{S}_4^{(2)} f_A^c f_B^c \mathcal{O}_{12} + S_9 S_{10} f_A^c f_B^c \mathcal{O}_{22}. \end{aligned} \tag{C24}$$

APPENDIX D: GAUGE COUPLING EVOLUTION

In this appendix, we derive the GQW equations for gauge coupling unification in the $\mathbb{Z}_6$ string models. We shall work in the orbifold GUT limit and also make some simplification assumptions of the matter spectra.

It is well known that gauge coupling unification in heterotic string models has a serious problem in the perturbative regime [34, 38]. From a simple reduction of the 10d string effective action, one finds $G_N = \mathrm{e}^{2\phi}\alpha'^4/128\pi V$, $\alpha_{\text{string}} = \mathrm{e}^{2\phi}\alpha'^3/16\pi V$, where ϕ is the dilaton, V the volume of the compactified space, and α' the Regge slope, thus[19]

$$M_{\text{Pl}} = \sqrt{\frac{8}{\alpha_{\text{string}}}} M_{\text{string}}. \tag{D1}$$

[19] A more careful definition, taking into account of the renormalization scheme dependence, is given in ref. [31].

Assuming $V \sim M_{\mathrm{GUT}}^{-6}$ (where $M_{\mathrm{GUT}} \simeq 3 \times 10^{16}$ GeV), $\alpha_{\mathrm{string}} = \alpha_{\mathrm{GUT}} \simeq 1/24$ and $\mathrm{e}^{2\phi} < 1$, one gets $M_{\mathrm{Pl}} \leq \alpha_{\mathrm{GUT}}^{-2/3} M_{\mathrm{GUT}} \simeq 2 \times 10^{17}$ GeV, which is apparently incorrect. Note that this result relies on the assumptions that the compactified space is symmetric, i.e., all the dimensions are of similar sizes, and there are no additional states near the GUT scale. Both assumptions are invalid in our models.

In heterotic models, it is well known that gauge couplings can obtain potentially large threshold corrections only if the models contain $\mathcal{N} = 2$ sub-sectors [32, 36]. Our $\mathbb{Z}_6$ orbifold models have exactly this kind of property, since both the $\mathbb{Z}_2$ and $\mathbb{Z}_3$ sub-orbifold twists leave exactly one complex plane unrotated. It is, however, a very complicate matter to compute gauge threshold corrections in string theory in the presence of discrete Wilson lines. (The simplest way to compute these corrections is using the target space duality symmetry [57], however in the presence of discrete Wilson lines the duality groups are broken to their discrete subgroups [58]. Even for the much simpler $\mathbb{Z}_3$ models, threshold corrections are only known numerically [35].[20])

Fortunately we can use a much simpler field theoretical method to compute gauge threshold corrections in the orbifold GUT limit, i.e., when one of the SO_4 dimensions is much larger than the string length scale and other compactified dimensions.[21] In this limit, all the winding modes and KK modes in small dimensions have string scale mass, and according to refs. [32, 36] one would expect the dominant contributions to the threshold corrections come from the KK modes of the large dimension. From now on, we shall assume this is correct and neglect all the contributions from states with string scale mass. The gauge coupling at the string scale, α_{string}, imposes a boundary condition for the renormalization group equations at the cut-off scale, which is taken to be the string scale, M_{string}.

We follow the field theoretical analysis in ref. [61] (see also [62]). It has been shown there the correction to a generic gauge coupling due to a tower of KK states with masses $M_{\mathrm{KK}} = m/R$ is

$$\alpha^{-1}(\Lambda) = \alpha^{-1}(\mu_0) - \frac{b}{4\pi} \int_{r\Lambda^{-2}}^{r\mu_0^{-2}} \frac{\mathrm{d}t}{t}\, \theta_3\left(\frac{\mathrm{i}t}{\pi R^2}\right) , \tag{D2}$$

where the integration is over the Schwinger parameter t, μ_0 and Λ are the IR and UV cut-offs, and $r = \pi/4$ is a numerical factor. θ_3 is the Jacobi theta function, $\theta_3(t) = \sum_{m=-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\pi m^2 t}$, representing the summation over KK states.

[20] Threshold corrections in models with continuous Wilson lines [63] have been computed in refs. [33]. Our models with discrete Wilson lines, however, are not the limiting cases of those with continuous lines.

[21] Admittedly the field theoretical calculation suffers from the usual UV divergence. The result is sensitive to the cutoff scale and needs be dealt with caution. However, we do not expect the RG evolution of the difference of gauge couplings to be affected much by our field theoretical treatment.

In our models, there are several modifications in the calculation. Firstly there are four sets of KK towers, with mass $M_{\rm KK} = m/R$ (for $P = P' = +$), $(m+1)/R$ (for $P = P' = -$) and $(m+1/2)/R$ (for $P = +$, $P' = -$ and $P = -$, $P' = +$), where $m \geq 0$. The summations over KK states give respectively $\frac{1}{2}\left(\theta_3(\mathrm{i}t/\pi R^2) - 1\right)$ for the first two cases and $\frac{1}{2}\theta_2(\mathrm{i}t/\pi R^2)$ for the last two (where $\theta_2(t) = \sum_{m=-\infty}^{\infty} \mathrm{e}^{\mathrm{i}\pi(m+1/2)^2 t}$), and we have separated out the zero modes in the $P = P' = +$ case. Secondly, the PS symmetry in our models must be further broken down to that of the SM. This breaking, in principle, can be induced by brane or bulk states. However, in sect. IV B we have shown this breaking is more likely due to non-renormalizable couplings of the states in the $T_{2,4}$ sectors. Since these states are identified with bulk states in the orbifold GUT limit (c.f. sect. III A), we shall assume, in what follows, the breaking of PS to the SM is bulk breaking. We further assume the breaking scale, $M_{\rm PS}$, is smaller than or equal to the compactification scale, M_c, so that we can neglect mass corrections to the massive KK states.

Tracing the renormalization group evolution from low energy scales, we are first in the realm of the MSSM, and the beta function coefficients are $b_i^{\rm MSSM} = (\frac{33}{5}, 1, -3)$. When we pass the PS breaking scale, $M_{\rm PS}$, the beta function coefficients become $(b_{++}^{\rm PS} + b_{\rm brane})_i$, where the two terms represent contributions from the bulk and brane states. These coefficients can contain contributions from additional states besides those of the MSSM, for example, from a vector-like pair $(\mathbf{4},\mathbf{1},\mathbf{1}) + (\overline{\mathbf{4}},\mathbf{1},\mathbf{1})$. The next energy threshold is the compactification scale M_c. From this scale to the string scale, we have the four sets of KK states. Since we have assumed $M_{\rm PS} \ll M_c$, the beta function coefficients are those of the complete PS representations.

Collecting these facts, and using $\theta_2(\mathrm{i}t/\pi R^2) \simeq \theta_3(\mathrm{i}t/\pi R^2) \simeq \sqrt{\frac{\pi}{t}}R$ for $t/R^2 \ll 1$, we find the GQW equations,

$$\begin{aligned}\frac{2\pi}{\alpha_i(\mu)} \;\simeq\; & \frac{2\pi}{\alpha_{\rm string}} + b_i^{\rm MSSM}\log\frac{M_{\rm PS}}{\mu} + (b_{++}^{\rm PS} + b_{\rm brane})_i \log\frac{M_{\rm string}}{M_{\rm PS}} \\ & - \frac{1}{2}(b_{++}^{\rm PS} + b_{--}^{\rm PS})_i \log\frac{M_{\rm string}}{M_c} + b^{\mathcal{G}}\left(\frac{M_{\rm string}}{M_c} - 1\right), \end{aligned} \tag{D3}$$

for $i = 1, 2, 3$, where we have taken the cut-off scales, $\mu_0 = M_c$ and $\Lambda = M_{\rm string}$. $b^{\mathcal{G}} = \sum_{P=\pm,P'=\pm} b_{PP'}^{\rm PS}$, so in fact it is the beta function coefficient of the orbifold GUT gauge group, $\mathcal{G}$. The beta function coefficients in the last two terms have an $\mathcal{N} = 2$ nature, since the massive KK states enjoy a large supersymmetry. For $\mathcal{G} = \mathrm{SO}_{10}$ and E_6, $b^{\mathcal{G}} = -16 + 4n_{\mathbf{16}} + 2n_{\mathbf{10}}$ and $-24 + 6n_{\mathbf{27}}$, where $n_{\mathbf{10}}$, $n_{\mathbf{16}}$ and $n_{\mathbf{27}}$ are numbers of bulk hypermultiplets in the respective representations. To accomplish bulk breaking, we need $n_{\mathbf{27}} = 4$ in models A1/B and $n_{\mathbf{16}} = 4$, $2 \leq n_{\mathbf{10}} \leq 6$ in model A2, therefore $b^{\mathrm{E}_6} = 0$ and $b^{\mathrm{SO}(10)} = 2n_{\mathbf{10}}$. Coincidentally there is no power-law running in the E_6

model.

---

[1] M. B. Green, J. H. Schwarz and E. Witten, *Superstring theory*, Cambridge University Press (1987);
D. Lüst and S. Theisen, *Lectures on string theory*, Springer-Verlag (1989);
J. Polchinski, *String theory*, Cambridge University Press (1998).

[2] D. J. Gross, J. A. Harvey, E. J. Martinec and R. Rohm, Phys. Rev. Lett. **54**, 502 (1985);
D. J. Gross, J. A. Harvey, E. J. Martinec and R. Rohm, Nucl. Phys. B **256**, 253 (1985);
D. J. Gross, J. A. Harvey, E. J. Martinec and R. Rohm, Nucl. Phys. B **267**, 75 (1986).

[3] P. Candelas, G. T. Horowitz, A. Strominger and E. Witten, Nucl. Phy. B **258**, 46 (1985).

[4] L. J. Dixon, J. A. Harvey, C. Vafa and E. Witten, Nucl. Phys. B **261**, 678 (1985);
L. J. Dixon, J. A. Harvey, C. Vafa and E. Witten, Nucl. Phys. B **274**, 285 (1986).

[5] B. Schellekens (ed), *Superstring construction*, North-Holland (1989).

[6] D. Bailin and A. Love, Phys. Rept. **315**, 285 (1999).

[7] For reviews, see e.g.,
C. Angelantonj and A. Sagnotti, Phys. Rept. **371**, 1 (2002) [Erratum-ibid. **376**, 339 (2003)];
A. M. Uranga, Class. Quant. Grav. **20**, S373 (2004);
D. Lüst, Class. Quant. Grav. **21**, S1399 (2004).

[8] B. R. Greene, K. H. Kirklin, P. J. Miron and G. G. Ross, Nucl. Phys. B **278**, 667 (1986);
B. R. Greene, K. H. Kirklin, P. J. Miron and G. G. Ross, Nucl. Phys. B **292**, 606 (1987).

[9] L. E. Ibáñez, J. E. Kim, H.-P. Nilles and F. Quevedo, Phys. Lett. B **191**, 282 (1987);
D. Bailin, A. Love and S. Thomas, Phys. Lett. B **194**, 385 (1987);
J. A. Casas and C. Muñoz, Phys. Lett. B **209**, 214 (1988);
J. A. Casas, M. Mondragon and C. Muñoz, Phys. Lett. B **230**, 63 (1989).

[10] L. E. Ibáñez, J. Mas, H. P. Nilles and F. Quevedo, Nucl. Phys. B **301**, 157 (1988).

[11] A. Font, L. E. Ibáñez, F. Quevedo and A. Sierra, Nucl. Phys. B **331**, 421 (1990).

[12] B. A. Campbell, J. R. Ellis, J. S. Hagelin, D. V. Nanopoulos and R. Ticciati, Phys. Lett. B **198**, 200 (1987);
I. Antoniadis, J. R. Ellis, J. S. Hagelin and D. V. Nanopoulos, Phys. Lett. B **208**, 209 (1988) [Addendum-ibid. B **213**, 562 (1988)];
I. Antoniadis, J. R. Ellis, J. S. Hagelin and D. V. Nanopoulos, Phys. Lett. B **231**, 65 (1989);
A. E. Faraggi, D. V. Nanopoulos and K. J. Yuan, Nucl. Phys. B **335**, 347 (1990);
A. E. Faraggi, Phys. Lett. B **326**, 62 (1994);
G. B. Cleaver, A. E. Faraggi and D. V. Nanopoulos, Phys. Lett. B **455**, 135 (1999);
G. B. Cleaver, A. E. Faraggi and D. V. Nanopoulos, Int. J. Mod. Phys. A **16**, 425 (2001).

[13] G. Aldazabal, A. Font, L. E. Ibáñez and A. M. Uranga, Nucl. Phys. B **452**, 3 (1995);

G. Aldazabal, A. Font, L. E. Ibáñez and A. M. Uranga, Nucl. Phys. B **465**, 34 (1996);

S. Chaudhuri, G. Hockney and J. Lykken, Nucl. Phys. B **469**, 357 (1996);

Z. Kakushadze and S. H. H. Tye, Phys. Rev. Lett. **77**, 2612 (1996);

Z. Kakushadze and S. H. H. Tye, Phys. Rev. D **54**, 7520 (1996).

[14] I. Antoniadis, G. K. Leontaris and J. Rizos, Phys. Lett. B **245**, 161 (1990);

G. K. Leontaris and N. D. Tracas, Phys. Lett. B **372**, 219 (1996).

[15] Some attempts on flavor physics in heterotic models are:

L. E. Ibáñez, Phys. Lett. B **181**, 269 (1986);

J. Casas and C. Muñoz, Nucl. Phys. B **332**, 189 (1990), [Erratum-ibid, B **340**, 280 (1990)];

T. T. Burwick, R. K. Kaiser and H. F. Muller, Nucl. Phys. B **355**, 689 (1991);

J. A. Casas, F. Gomez and C. Muñoz, Phys. Lett. B **292**, 42 (1992);

J. A. Casas, F. Gomez and C. Munoz, Int. J. Mod. Phys. A **8**, 455 (1993);

T. Kobayashi, Phys. Lett. B **354**, 264 (1995);

T. Kobayashi and O. Lebedev, Phys. Lett. B **566**, 164 (2003);

P. Ko, T. Kobayashi and J. h. Park, hep-ph/0406041.

[16] M. Dine and N. Seiberg, Phys. Lett. B **162**, 299 (1985).

[17] G. L. Cardoso, G. Curio, G. Dall'Agata, D. Lust, P. Manousselis and G. Zoupanos, Nucl. Phys. B **652**, 5 (2003);

K. Becker, M. Becker, K. Dasgupta and P. S. Green, JHEP **0304**, 007 (2003);

K. Becker, M. Becker, K. Dasgupta, P. S. Green and E. Sharpe, Nucl. Phys. B **678**, 19 (2004).

[18] S. Dimopoulos, S. Raby and F. Wilczek, Phys. Rev. D **24**, 1681 (1981);

J. Ellis, S. Kelly and D. V. Nanopoulos, Phys. Lett. B **249**, 441 (1990);

C. Giunti, C. W. Kim and U. W. Lee, Mod. Phys. Lett. A **6**, 1745 (1991);

U. Amaldi, W. de Boer and H. Fürstenau, Phys. Lett. B **260**, 447 (1991);

P. Langacker and M. X. Luo, Phys. Rev. D **44**, 817 (1991).

[19] Y. Kawamura, Prog. Theor. Phys. **103**, 613 (2000);

G. Altarelli and F. Feruglio, Phys. Lett. B **511**, 257 (2001);

A. Hebecker and J. March-Russell, Nucl. Phys. B **613**, 3 (2001);

T. Asaka, W. Buchmüller and L. Covi, Phys. Lett. B **523**, 199 (2001);

L. Hall and Y. Nomura, Phys. Rev. D **64**, 055003 (2001);

L. Hall and Y. Nomura, Phys. Rev. D **66**, 075004 (2002);

R. Barbieri, L. J. Hall and Y. Nomura, Nucl. Phys. B **624**, 63 (2002);

L. J. Hall, H. Murayama and Y. Nomura, Nucl. Phys. B **645**, 85 (2002);

L. J. Hall, Y. Nomura, T. Okui and D. R. Smith, Phys. Rev. D **65**, 035008 (2002).

[20] R. Dermíšek and A. Mafi, Phys. Rev. D **65**, 055002 (2002);

H. D. Kim and S. Raby, JHEP **0301**, 056 (2003).

[21] T. Kobayashi, S. Raby and R.-J. Zhang, hep-ph/0403065, Phys. Lett. B **593**, 262 (2004).

[22] S. Forste, H. P. Nilles, P. K. S. Vaudrevange and A. Wingerter, hep-th/0406208.

[23] J. C. Pati and A. Salam, Phys. Rev. D **8**, 1240 (1973);
R. E. Marshak and R. N. Mohapatra, Phys. Lett. B **91**, 222 (1980).

[24] Y. Katsuki, Y. Kawamura, T. Kobayashi, N. Ohtsubo, Y. Ono and K. Tanioka, Nucl. Phys. B **341**, 611 (1990);
Y. Katsuki, Y. Kawamura, T. Kobayashi, N. Ohtsubo, Y. Ono and K. Tanioka, *Tables of Z_N orbifold models*, preprint DPKU-8904, Kanazawa (1989).

[25] L. E. Ibáñez, H. P. Nilles and F. Quevedo, Phys. Lett. B **187**, 25 (1987).

[26] C. Vafa, Nucl. Phys. B **273**, 592 (1986);
D. Freed and C. Vafa, Commun. Math. Phys. **110**, 340 (1987).

[27] H. Georgi, H. R. Quinn and S. Weinberg, Phys. Rev. Lett. **33**, 451 (1974).

[28] F. Gliozzi, J. Scherk and D. I. Olive, Nucl. Phys. B **122**, 253 (1977).

[29] A. N. Schellekens and N. P. Warner, Phys. Lett. B **181**, 339 (1986);
A. N. Schellekens and N. P. Warner, Nucl. Phys. B **287**, 317 (1987);
W. Lerche, A. N. Schellekens and N. P. Warner, Nucl. Phys. B **299**, 91 (1988).

[30] S. Dimopoulos and F. Wilczek, in *The unity of the fundamental interactions*, ed. A. Zichichi, Plenum (1983);
A. Masiero, D. V. Nanopoulos, K. Tamvakis and T. Yanagida, Phys. Lett. B **115**, 380 (1982);
B. Grinstein, Nucl. Phys. B **206**, 387 (1982).

[31] V. S. Kaplunovsky, Nucl. Phys. B **307**, 145 (1988) [Erratum-ibid. B **382**, 436 (1992)].

[32] L. Dixon, V. S. Kaplunovsky and J. Louis, Nucl. Phys. B **355**, 649 (1991);
P. Mayr and S. Stieberger, Nucl. Phys. B **407**, 725 (1993).

[33] P. Mayr and S. Stieberger, Phys. Lett. B **355**, 107 (1995);
T. Kawai, Phys. Lett. B **372**, 59 (1996);
J. A. Harvey and G. W. Moore, Nucl. Phys. B **463**, 315 (1996);
M. Henningson and G. W. Moore, Nucl. Phys. B **482**, 187 (1996);
S. Stieberger, Nucl. Phys. B **541**, 109 (1999).

[34] K. Dienes, Phys. Rept. **287**, 447 (1997).

[35] P. Mayr, H.-P. Nilles, and S. Stieberger, Phys. Lett. B **317**, 53 (1993).

[36] V. S. Kaplunovsky and J. Louis, Nucl. Phys. B **444**, 191 (1995);
B. de Wit, V. Kaplunovsky, J. Louis and D. Lüst, Nucl. Phys. B **451**, 53 (1995);
H. P. Nilles and S. Stieberger, Nucl. Phys. B **499**, 3 (1997).

[37] P. Hořava and E. Witten, Nucl. Phys. B **460**, 506 (1996);
P. Hořava and E. Witten, Nucl. Phys. B **475**, 94 (1996).

[38] E. Witten, Nucl. Phys. B **471**, 135 (1996);
T. Banks and M. Dine, Nucl. Phys. B **479**, 173 (1996).

[39] V. Lucas and S. Raby, Phys. Rev. D **55**, 6986 (1997);

H. Murayama and A. Pierce, Phys. Rev. D **65**, 055009 (2002).

[40] N. Tsutsui *et al.* [CP-PACS Collaboration], hep-lat/0402026;
S. Aoki *et al.* [JLQCD Collaboration], Phys. Rev. D **62**, 014506 (2000);
Y. Aoki [RBC Collaboration], Nucl. Phys. Proc. Suppl. **119**, 380 (2003).

[41] C. K. Jung, Brookhaven National Lab preprint, UNO 02-BNL (2002).

[42] M. B. Green and J. H. Schwarz, Phys. Lett. B **149**, 117 (1984).

[43] E. Witten, Phys. Lett. B **149**, 351 (1984);
M. Dine, N. Seiberg and E. Witten, Nucl. Phys. B **289**, 589 (1987);
W. Lerche, B. Nilsson and A. N. Schellekens, Nucl. Phys. B **289**, 609 (1987);
J. J. Atick, L. J. Dixon and A. Sen, Nucl. Phys. B **292**, 109 (1987);
T. Kobayashi and H. Nakano, Nucl. Phys. B **496**, 103 (1997).

[44] J. A. Casas, E. K. Katehou and C. Muñoz, Nucl. Phys. B **317**, 171 (1989);
G. Cleaver, M. Cvetič, J. R. Espinosa, L. L. Everett and P. Langacker, Nucl. Phys. B **525**, 3 (1998).

[45] B. C. Allanach, S. F. King, G. K. Leontaris and S. Lola, Phys. Rev. D **56**, 2632 (1997).

[46] K. Narain, Phys. Lett. B **169**, 41 (1986);
K. Narain, M. Sarmadi and E. Witten, Nucl. Phys. B **279**, 369 (1987).

[47] T. Kobayashi and N. Ohtsubo, Phys. Lett. B **257**, 56 (1991).

[48] T. Kobayashi and N. Ohtsubo, Int. J. Mod. Phys. A **9**, 87 (1994).

[49] T. Kobayashi and N. Ohtsubo, Phys. Lett. B **245**, 441 (1990).

[50] J. Giedt, Annals Phys. **297**, 67 (2002).

[51] J. E. Kim, Phys. Lett. B **564**, 35 (2003).

[52] H.-P. Nilles, talk at the String Phenomenology '04, Ann Arbor, Michigan.

[53] D. C. Lewellen, Nucl. Phys. B **337**, 61 (1990);
A. Font, L. E. Ibáñez and F. Quevedo, Nucl. Phys. B **345**, 389 (1990).

[54] D. Friedan, E. Martinec and S. Shenker, Phys. Lett. B **160**, 55 (1985);
D. Friedan, E. Martinec and S. Shenker, Nucl. Phys. B **271**, 93 (1986);
V. G. Knizhnik, Phys. Lett. B **160**, 403 (1985).

[55] S. Hamidi and C. Vafa, Nucl. Phys. B **279**, 465 (1987);
L. Dixon, D. Friedan, E. Martinec, and S. Shenker, Nucl. Phys. B **282**, 13 (1987).

[56] J. P. Derendinger, S. Ferrara, C. Kounnas and F. Zwirner, Nucl. Phys. B **372**, 145 (1992);
L. E. Ibáñez and D. Lüst, Nucl. Phys. B **382**, 305 (1992);
V. S. Kaplunovsky and J. Louis, Phys. Lett. B **306**, 269 (1993);
A. Brignole, L. Ibáñez and C. Muñoz, Nucl. Phys. B **422**, 125 (1994) [Erratum-ibid. B **436**, 747 (1995)];
A. Brignole, L. E. Ibáñez, C. Muñoz and C. Scheich, Z. Phys. C **74**, 157 (1997).

[57] A. Giveon, M. Porrati and E. Rabinovici, Phys. Rept. **244**, 77 (1994).

[58] M. Spalinski, Phys. Lett. B **275**, 47 (1992);
J. Erler, D. Jungnickel, and H.-P. Nilles, Phys. Lett. B **276**, 303 (1992);

J. Erler and M. Spalinski, Int. J. Mod. Phys. A **9**, 4407 (1994);
D. Bailin, A. Love, W. A. Sabra and S. Thomas, Mod. Phys. Lett. A **9**, 1229 (1994).

[59] M. Dine, N. Seiberg, X. G. Wen and E. Witten, Nucl. Phys. B **278**, 769 (1986);
M. Dine, N. Seiberg, X. G. Wen and E. Witten, Nucl. Phys. B **289**, 319 (1987);
J. Distler and B. Greene, Nucl. Phys. B **304**, 1 (1988).

[60] A. Font, L. E. Ibanez, H. P. Nilles and F. Quevedo, Nucl. Phys. B **307**, 109 (1988);
A. Font, L. E. Ibanez, H. P. Nilles and F. Quevedo, Phys. Lett. **210B**, 101 (1988) [Erratum-ibid. B **213**, 564 (1988)].

[61] K. R. Dienes, E. Dudas and T. Gherghetta, Phys. Lett. B **436**, 55 (1998);
K. R. Dienes, E. Dudas and T. Gherghetta, Nucl. Phys. B **537**, 47 (1999).

[62] D. M. Ghilencea and S. Groot Nibbelink, Nucl. Phys. B **641**, 35 (2002);
D. M. Ghilencea, Nucl. Phys. B **670**, 183 (2003).

[63] T. Mohaupt, Int. J. Mod. Phys. A **9**, 4637 (1994);
G. Lopes Cardoso, D. Lüst and T. Mohaupt, Nucl. Phys. B **432**, 68 (1994).